

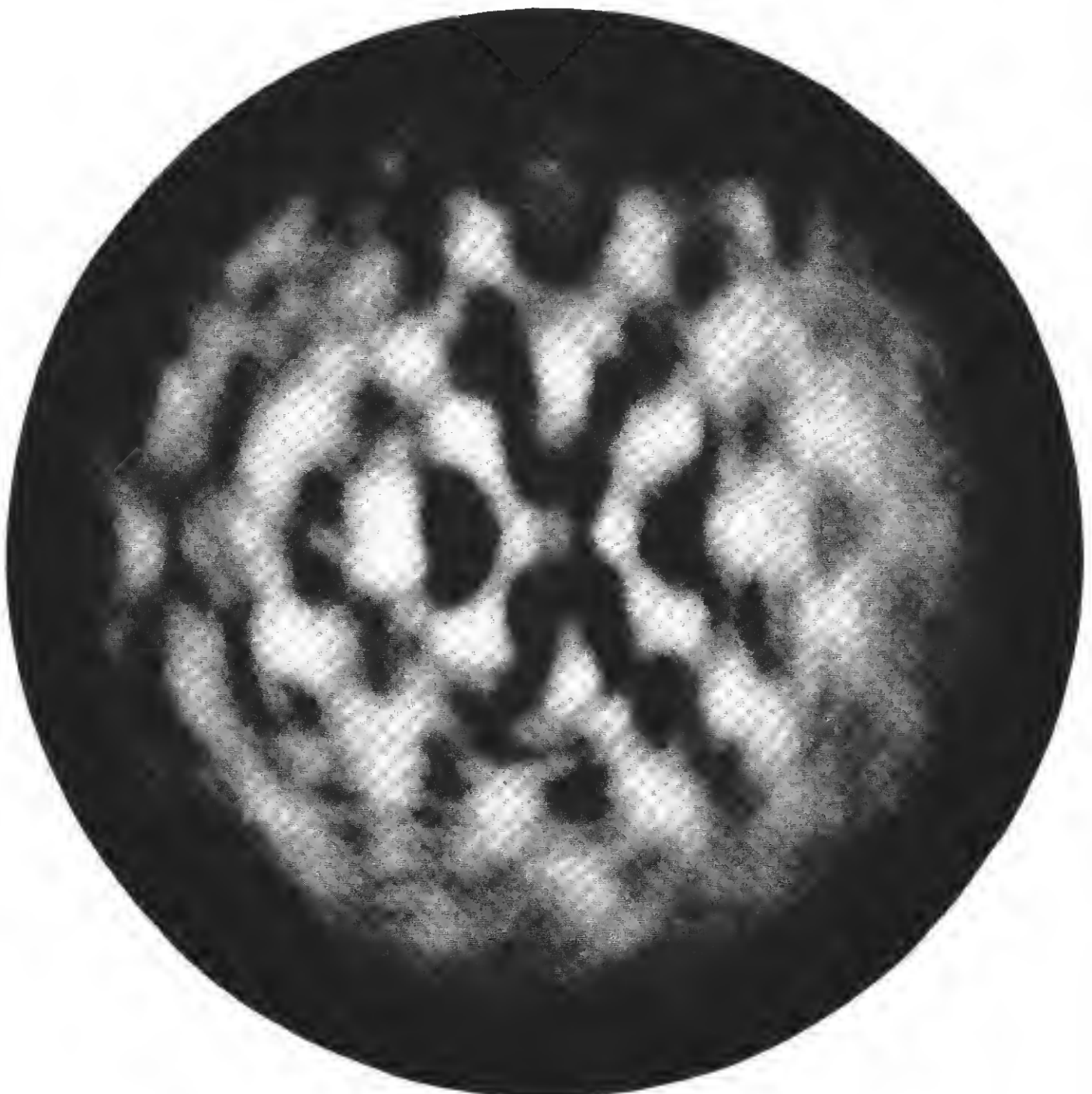
BBC ENGINEERING

D R PHILLIPS

D JSE
E RSS
31 June

Number 105

January 1977



BBC Engineering

including Engineering Division Monographs

A record of BBC technical experience and developments
in radio and television broadcasting

Published approximately four times per year by
BBC Publications,
35 Marylebone High Street,
London W1M 4AA
ISSN 0308—2369
ISBN 0 563 17294 0

Edited by
BBC Engineering Information Department,
Broadcasting House,
London W1A 1AA

Printed by
The Yale Press Ltd,
London SE25 5LY.

Price
80p or \$2 per issue post free

Annual subscription
£3 or \$8 for 4 issues post free.

Articles are copyright © 1977 by the
British Broadcasting Corporation
and may not be reproduced in whole
or part without written permission.

Contents

Editorial	2
Principal articles	
Holography: General Principles and Applications to Sound and Television Recording E. W. Taylor	4
Digital Recording using Hologram Arrays: Low-frequency Micro-holograms on Photographic Film K. Hacking	21

The cover illustration is a hologram (magnified about
4000 diameters) of a digit plate used in experimental work
described in the article by K. Hacking.

Editorial

A favourite topic for the 'back-room boys' is video recording. Magnetic recording is no longer a technical marvel but its successor is hard to find. Has it got significant deficiencies or is it just a happy hunting ground for 'boffins'?

There are three main areas where improvements are sought. First, running costs; i.e. staff, capital depreciation, and the cost of the recording medium. The recorder is now a production tool whose costs are not insignificant. Second, programme production often requires multi-generation copies and in any coded colour signal some degradation in technical quality occurs. Third, the cost of the stored tape is too high for it to be a good solution to the archival storage problem. Colour film is even more expensive and the colours are fugitive.

Existing analogue VTRs are difficult to depose for two reasons: the variables of the frequency modulation system have been adapted fully to utilize the recording medium so that little improvement in tape consumption is likely to occur; and the ability to erase and re-record is valuable. It is an important asset in studio productions although, in practice, the re-usable feature only saves, say, one third of the total tape requirement.

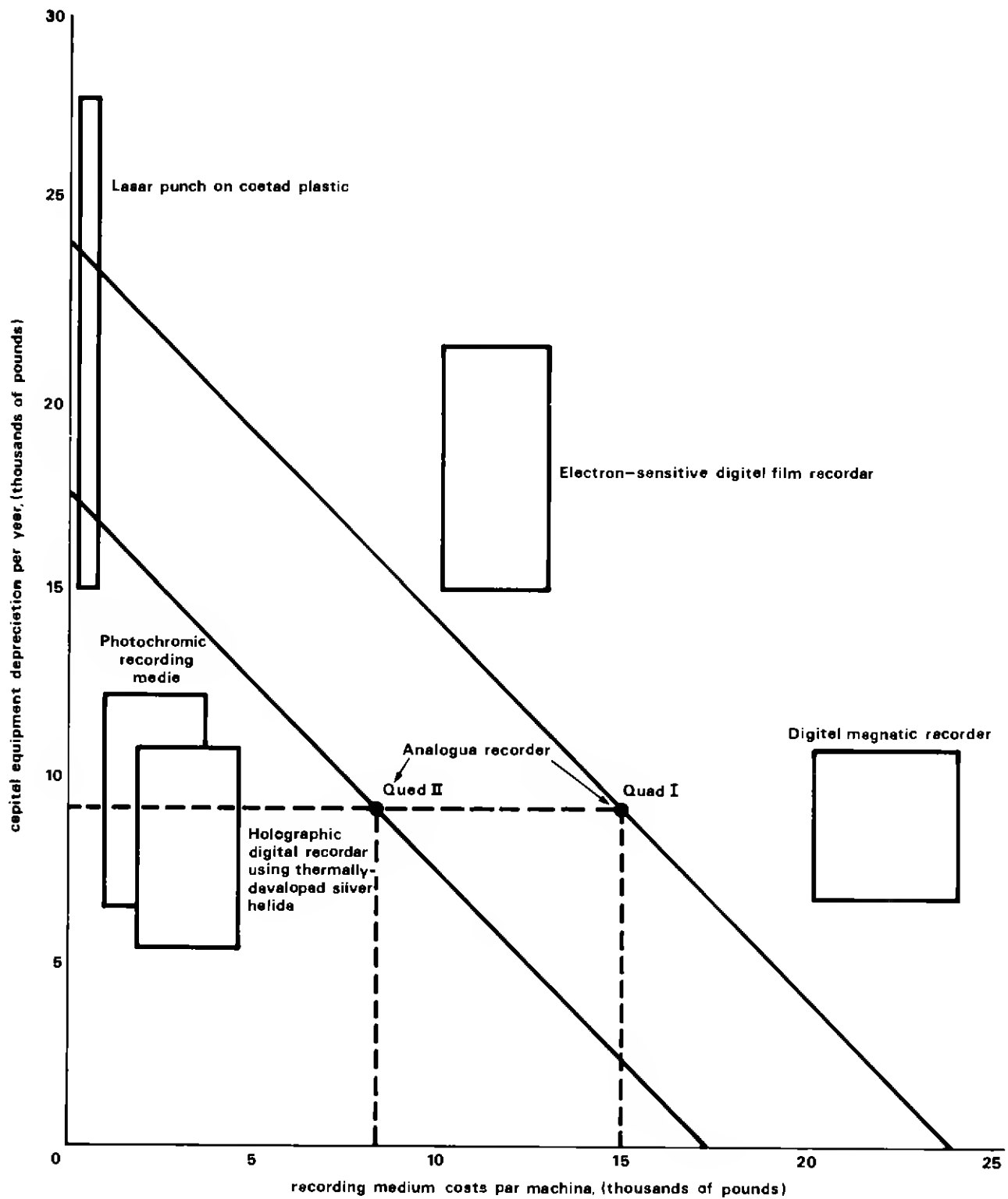
There seems to be a growing case for two types of machine: one which is used solely for time delay/archival storage purposes, and one dedicated to studio production applications. The former needs to have a low consumption

of recording medium but can be free from the complexities of studio production equipment. The latter will inevitably be a complicated machine with a high capital cost.

In most technical discussion, digital techniques are said to have the greatest potential to meet these needs. The accompanying diagram illustrates the position. Some techniques give lower running costs; others lower capital costs.

For archival storage the medium cost is of first importance. Even though the figures cannot be precise, some of the new techniques hold great promise of cost reductions. The application of these techniques to data recording opens up an additional market which is a spur to our endeavours. Work reported in this journal on holographic recording techniques could play a significant part in future developments.

The digital equivalent of the studio recorder is more difficult to justify. The ability to record without degradation makes any machine a valuable research tool but the tape consumption makes it unsuitable for normal studio use. We need a considerable step forward in bit-reduction coding techniques to reduce the tape consumption and a change in studio techniques fully to exploit the virtues of the digital recorder. Perhaps we shall see a convergence of sound and television techniques with the introduction of multitrack video and sound recorders.



Holography: General Principles and Applications to Sound and Television Recording

E. W. Taylor, MA (Cantab), CEng, MIEE

Summary: The history of the development of holography is surveyed and a theoretical introduction is given. The theoretical and practical aspects of using holographic techniques in both analogue and digital recording systems are examined, and comparisons are made between holographic and other present-day or proposed methods of recording. The use of holographic techniques in television broadcast systems is briefly discussed.

- 1 Introduction and historical survey
- 2 Principles of holography
 - 2.1 The surface hologram
 - 2.2 The volume hologram
- 3 Characteristics of transmission and recording systems
 - 3.1 General
 - 3.2 Summary of principles of digital transmission
 - 3.3 General comments on digital recording
- 4 The use of holographic techniques in recording systems
 - 4.1 General
 - 4.2 Analogue recording
 - 4.3 Digital recording
- 5 The maximum packing density of a holographic recording system
- 6 Practical considerations
 - 6.1 Optical arrangements
 - 6.2 Recording and playback
 - 6.3 Hologram size
 - 6.4 Present-day feasibility of implementing a holographic digital recording system
- 7 Comparison of holographic and other proposed digital recording systems
- 8 Recapitulation and conclusions
- 9 References

1 Introduction and historical survey

In 1949 Gabor^{1-3,4a} basing his ideas on principles suggested by Bragg,⁵ put forward a two-step method of optical imagery. In the first step the object (essentially a transparency) was illuminated with coherent monochromatic light and the pattern resulting from interference between light scattered from features in the object and light passing directly through it was recorded on a photographic plate. In the second step the processed plate itself was illuminated with monochromatic light. The diffraction of light by the pattern recorded on the plate gave rise to two additional components of radiation. If radiation of the same

wavelength and geometry was used in both halves of the two-step process, one of these components was identical to the original wave-front issuing from the object. This component corresponded to a 'reconstructed' image of the original object occupying the same position in space (relative to the plate) as the original object. The other component was the so-called twin-wave: this occurred because there were two object positions which could give rise to the same recorded interference pattern, and images in both these positions were reconstructed. The name 'hologram' was given to the recorded pattern, because (to quote from reference 2) "the photograph contains the total information required for reconstructing the object, which can be two-dimensional or three-dimensional".

Gabor's work was directed towards improving the resolving power of electron microscopes by forming the hologram of a small object in electron radiation and reconstructing the image of the object in light radiation using an appropriately enlarged version of the hologram. Because of technical difficulties concerned with electrical and mechanical stability this aspect of the work was not successful. Problems were also experienced in reconstructing the image: in particular, the presence of the twin-wave meant that the wanted image always had to be viewed against an unwanted out-of-focus background. The difficulty of obtaining sufficiently monochromatic and coherent light sources also contributed to the initial lack of interest in the subject.

A fairly simple method of eliminating the twin-wave problem was demonstrated by Leith and Upatnieks⁶ in 1962. This consisted of providing an off-axis 'reference wave', so that interference occurred between this wave and light passing through the object (which was still only considered as a transparency). The recorded interference pattern obtained with this arrangement had a regular grating-like structure and, as with a conventional diffraction grating, gave rise to two deviated components travelling in different directions when illuminated with monochromatic light. One of these components was the required reconstructed object wave-front and the other was the unwanted twin-wave: separation of the two was obtained simply by observing in the appropriate direction. Radiation from the same source must

be used to illuminate the object and at the same time provide the reference wave. This is relatively easy to arrange in the case of light radiation, by using a beam-splitting system, but in the case of electron radiation no beam-splitting system is available and consequently off-axis holography is not possible. Gabor recognised this limitation when stating² "it is very likely that in light optics, where beam splitters are available, methods can be found for providing the coherent background which will allow better separation of object planes and more effective elimination of the effects of the twin-wave than the simple arrangements which have been investigated".

Leith and Upatnieks were also responsible⁷ for introducing the idea of diffuse-illumination holography. This step was largely made possible by the development of the laser^{8,9} which made available, for the first time, a source of coherent light of adequate intensity and coherence length (by 'coherence length' is meant the greatest difference between two points along the direction of propagation of the radiation for which significant correlation exists between the phase of the radiation at each point). Several advantages were found in illuminating a transparent object with diffuse light. As the light passed through the object in many directions, light from any one part of the object was spread over the entire area of the hologram: thus any portion of the so-called 'diffuse hologram' formed in this way contained information concerning the entire object and a complete image can be reconstructed from it. Furthermore, surface imperfections on the hologram or other parts of the optical system (blemishes, dust particles, etc.) did not give rise to discrete effects in the reconstruction (as was the case with non-diffuse illumination), but only added to the overall flare or 'veiling glare' in the reconstructed image. (In fact, the effect of surface imperfections in conventional optical imaging systems is always greater when using non-diffuse illumination, and some degree of diffusion is used whenever possible.) This advantage is examined in greater detail in section 4.1. Another advantage of the use of diffuse illumination was that the whole reconstructed image could be viewed from a variety of directions without optical aids, whereas with non-diffuse illumination only the reconstruction of that part of the object lying on the line between the source of illumination and the eye could be so observed. In fact, if a three-dimensional diffusely-reflecting object was used instead of a transparency, the reconstructed (virtual) image was also three-dimensional and gave the appearance of observing the original object through a window formed by the hologram itself. The twin-wave reconstruction formed a real image which was also three-dimensional but which had 'reversed parallax' or 'pseudoscopic' properties.^{4b,7,10,11}

The holographic recording can be either two-dimensional (i.e. a cross-section of the interference pattern, giving rise to a 'surface' hologram) or three-dimensional, in which the thickness of the recording medium is large relative to the 'scale' (or average distance between corresponding features) of the interference pattern: this produces a 'volume' hologram. In the first case the twin-wave component is always reconstructed but in the second case it is theoretically possible to suppress it. This represents a considerable

advantage because, in principle, all the light incident on the hologram may be diverted into the wanted reconstructed beam^{12,13}. This aspect is discussed further in Section 2.2.

Diffraction of radiation will occur if point-to-point variations of either amplitude or phase are introduced. A hologram in which the recording of the original interference pattern takes the form of quasi-periodic spatial variations of transmission will introduce such amplitude changes into the reconstruction-beam radiation, giving rise to the diffracted image-forming components (wanted and twin-wave). Appropriate phase variations^{4c} will be introduced into the reconstruction-beam radiation if the hologram consists of transparent material of uniform thickness in which the refractive index at each point is related to the intensity of radiation in the interference pattern at that point during formation of the hologram. Alternatively, the refractive index of the material may remain constant and the phase changes be produced by variations in thickness of the material. Phase changes may also be produced by embossing or etching the diffraction pattern on a reflecting surface. There are a number of advantages in the use of such 'phase holograms' instead of the 'amplitude holograms' previously described. In the first place, the 'diffraction efficiency' can be greatly increased^{4d,12-17}. (The diffraction efficiency is the proportion of the incident radiant power which goes to form the wanted image). In addition, many substances can be used to control the relative phase of the incident radiation while comparatively few provide the necessary control of its amplitude; thus the range of materials available for use in holographic recording is greatly extended. In addition, light can penetrate a volume phase hologram of considerable thickness without undue attenuation, thus making the use of such holograms practicable.

It is worth digressing at this point to examine very briefly the possible use of the three-dimensional holographic image in a broadcast television system¹⁸. In place of a photographic plate or other form of recording medium, a photo-sensitive surface intercepts the holographic interference pattern, the intensity variations over the surface being converted into a time-varying electrical signal by a scanning system. This signal is suitably processed and transmitted. At the receiver the signal is supplied by way of another scanning system to a 'display' surface, having the property that the relative phases of the light transmitted through (or reflected by) small areas of the surface, or alternatively the optical transmission or reflectivity of the small areas, is related to the magnitude of the electrical signal supplied to the surface by the scanning system. A phase or amplitude hologram is thus built up on the surface which, when illuminated by a suitable reconstruction beam, will provide a holographic image of the original object.

When assessing new television transmission systems, the picture quality achieved by present-day systems must be taken as a standard of comparison. Any potential improvements in particular aspects of picture quality that may be obtained with the new system must represent a worthwhile gain in the system considered as a whole, and must not be accompanied by reductions in other aspects of picture quality that are currently taken as acceptable. Such con-

siderations enable the distinction to be made between a viable new system of television transmission on the one hand, and an interesting but nevertheless unacceptable 'laboratory curiosity' on the other. In the present case, the improvement in picture quality represented by the addition of three-dimensional information must not be made at the expense of reductions of standards in terms of screen size, portrayal of motion, colour and resolution. The first conclusion that is reached if these standards of judgement are adopted is that an extremely wide bandwidth video signal (of the order of 3×10^6 MHz) would be required. It would be totally impracticable in the foreseeable future to radiate such a signal, although it might possibly be conveyed by cable (conventional or fibre-optic) where a multi-channel approach could be used. In addition, devices for use in place of the conventional camera-tube and display, which are capable of handling either this high video bandwidth or the very fine holographic patterns, do not seem likely to be developed on any useful time-scale. Even if such components were to become available, difficulties would still exist in the provision of coherent scene lighting (since the difference in the source-to-receptor path lengths corresponding to different parts of the scene must not be greater than the coherence length of the radiation, which is at best of the order of a few metres), the exclusion of daylight or conventional artificial light, and the reproduction of scene colour. Thus although attractive and simple in principle, it has regrettably to be concluded that a three-dimensional television system of broadcast quality, using holographic techniques, is not feasible at the present time or in the foreseeable future.

2 Principles of holography

2.1 The surface hologram

Following the treatment given to the subject by Smith,^{4c} the theory of holography is introduced by a simple example

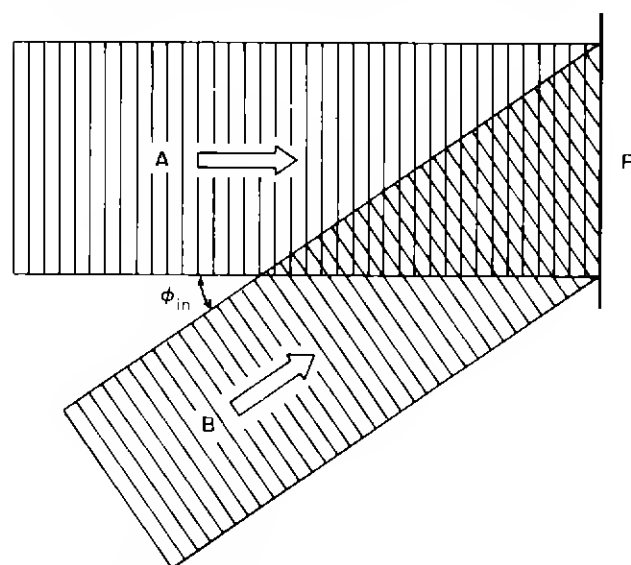


Fig. 1 Two interfering beams

which serves to illustrate the main features of the subject. Consider two collimated beams (A and B) of monochromatic and mutually coherent light travelling horizontally (and therefore containing vertical plane wavefronts) and falling on the same area of a vertical flat photographic plate (P in figure 1). This plate is taken to consist of a transparent substrate supporting a very thin photo-sensitive layer, on to which the incident light falls directly. For simplicity, one beam is assumed to be normally incident onto the plate while the angle of incidence of the other beam in a horizontal plane is ϕ_{in} . An interference pattern will be formed in the volume common to both light beams and the emulsion of the photographic plate will intercept a section of this pattern which will consist of vertical equally-spaced straight lines. Figure 2 shows a horizontal section of an

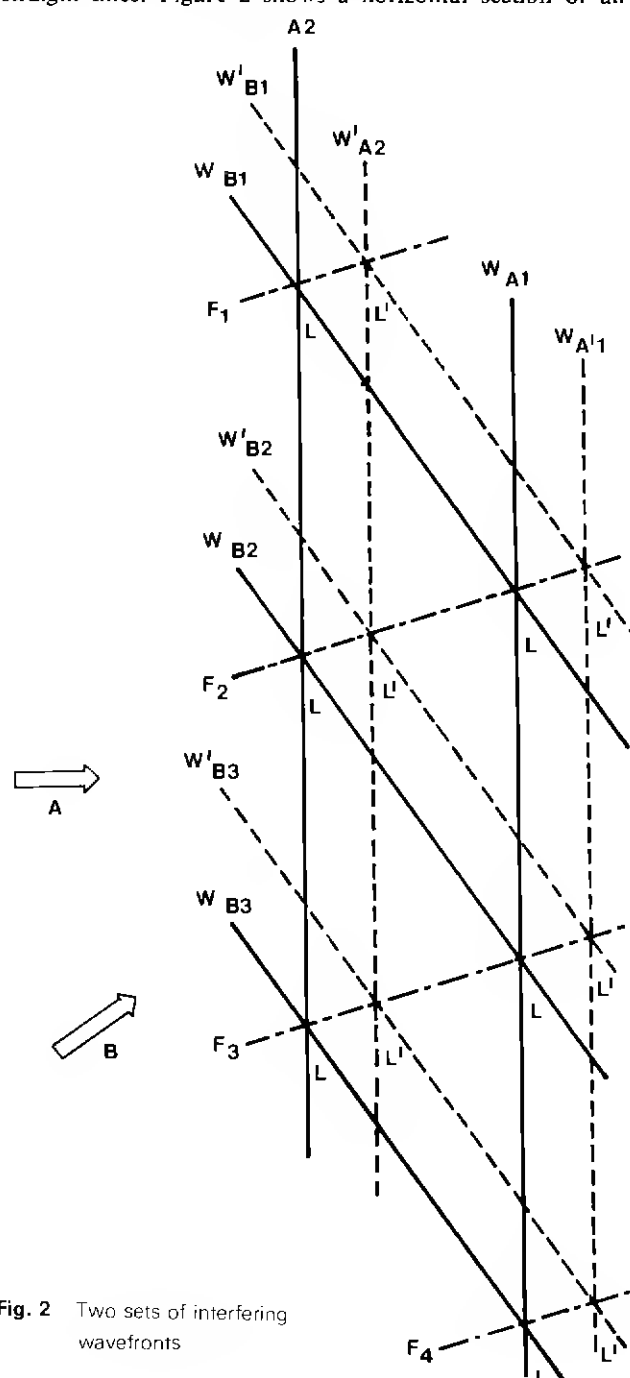


Fig. 2 Two sets of interfering wavefronts

enlargement of a region close to the photographic plate. If the wavelength of the radiation is λ , identical features in the electromagnetic field will occur in planes separated by this distance. W_{A1} and W_{A2} show the positions of two such planes, at a certain instant of time, in the beam A of figure 1, while W_{B1} , W_{B2} , etc. show corresponding features in the beam B. At the lines of intersection (L) of these two sets of planes the fields will reinforce each other to the greatest extent, and the maximum electromagnetic disturbance will occur. A short time later these features in the electromagnetic field will have moved to the planes $W_{A1'}$, etc. and $W_{B1'}$, etc. which intersect at the lines L' . The parallel planes F_1 , F_2 etc. containing the lines L , L' , etc. as the waves progress will correspond with the regions of greatest electromagnetic disturbance in the interference pattern: in optical terms, these regions will show as bright areas. These planes are stationary in space and the interference pattern is therefore stationary.

Consider the exposure of the photographic plate (P in figure 3) which is placed normal to the beam A, and there-

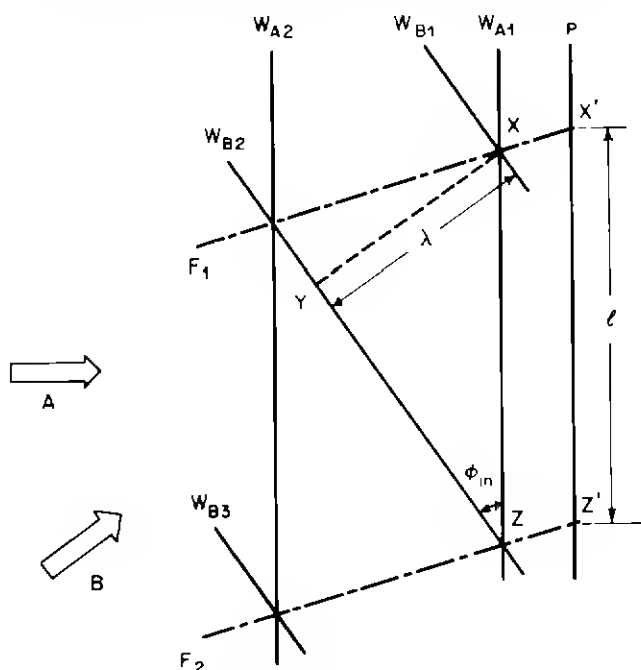


Fig. 3 Geometry of two-dimensional record of interference pattern

fore parallel to the wave-fronts W_{A1} , W_{A2} , etc. As the interference pattern is stationary, a time exposure of the plate may be made. Considering the right-angled triangle XYZ, it can be seen that

$$XZ = \frac{\lambda}{\sin \phi_{in}}$$

hence, because the planes F_1 and F_2 are parallel

$$X'Z' = l = \frac{\lambda}{\sin \phi_{in}} \quad (1)$$

The distance l is the spacing of the fringes on the photographic plate. When processed, the plate will carry a series

of alternately clear and opaque parallel stripes having the same spacing l , and will therefore behave as a diffraction grating. If illuminated by the beam A in figure 1, a number of beams will emerge from the plate at angles ϕ_n to the normal, where

$$\sin \phi_n = \pm n\lambda/l \quad (n = 0, 1, 2 \dots)$$

The condition $n = 0$ corresponds to the 'zero-order' or undeviated beam (A' in figure 4). The first-order ($n = 1$) diffracted beams (B' and B'' in figure 4) emerge at angles $\pm \phi_1$, where

$$\sin \phi_1 = \lambda/l \quad (2)$$

Comparing equations 1 and 2 it can be seen that $\phi_1 = \phi_{in}$ and that the Beam B' is therefore travelling in the same direction as the original beam B in figure 1. The beam B' may therefore be regarded as a 'reconstruction' of the beam B. The beam B'' corresponds to the unwanted twin-wave (see section 1). It may be regarded as the reconstruction of a hypothetical beam incident onto the plate at the same angle as the beam B in figure 1, but on the opposite side of beam A, since such a beam would give rise to the same interference pattern, and therefore the same stripe pattern on the processed plate, as the actual beam B.

If several 'object' beams (coming, for example, from different parts of an extended object) are simultaneously incident at various angles onto the plate, in addition to the 'reference' beam A, interference patterns will be formed between each such beam and the reference beam. On illuminating the processed plate, reconstructions of each such beam will be produced, together with the respective twin-waves. The complex interference pattern recorded on the plate may be regarded (following the treatment first proposed by Leith and Upatnieks⁶ and described by Smith^{4f}) as being composed of a number of spatial Fourier components. Each such component acts as a diffraction grating and diffracts radiation into two first-order components. The superposition of the components of radiation from all the spatial Fourier components gives the total radiation field diffracted

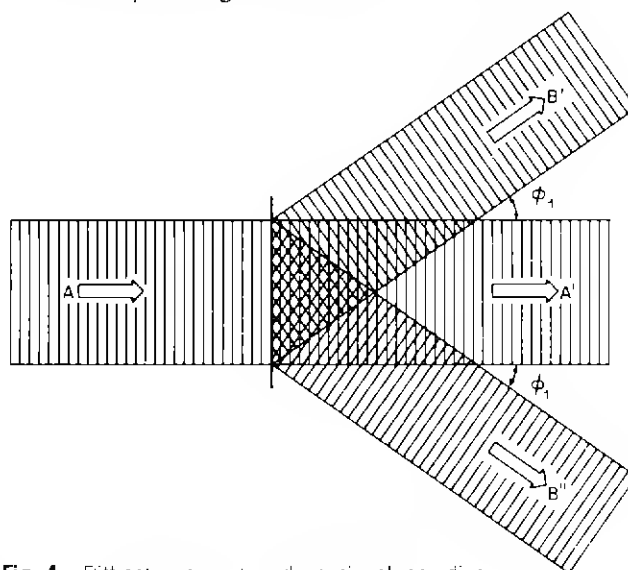


Fig. 4 Diffraction from two-dimensional recording

by the hologram and it can be shown that this field consists of a reconstruction of the original object wave, together with a conjugate twin-wave, as previously described. A plane wave arriving at the photographic plate along the direction of propagation of the reference wave, and another arriving along a direction related to the overall position of the object, would give rise to a spatial Fourier component having a certain periodicity and orientation. This Fourier component may be regarded as a spatial 'carrier', all the other Fourier components of the recorded diffraction pattern being regarded as 'sidebands' carrying information about the characteristics of the object and the reference beam. As the angle between the reference beam and object directions is increased, the spatial frequency of the carrier is increased, but so too is the angle of diffraction of the reconstructed radiation. In a practical holographic system a compromise must always be made between these two factors: a larger angle of diffraction assists in the separation of the wanted and unwanted diffracted components but demands a higher resolution from the recording medium. Figure 5 shows an enlargement of a portion of a hologram. It can be seen that individual elements of the recorded interference pattern (the fine structure in figure 5: see below) have a general similarity in size and orientation, giving the hologram a pronounced uniformity of texture: this corresponds to the presence of the spatial carrier. On the other hand, individual elements of the recorded pattern differ greatly in their opacity, as well as to some extent in their size and orientation. These differences correspond to the spatial sidebands and carry the object information.

If coherent light is reflected from a diffuse surface, mutual interference between components of reflected light emanating from different elementary portions of the surface

takes place. This gives rise to a random granular pattern, usually known as the 'speckle pattern', which appears superimposed on the object when it is observed visually and is recorded on a hologram of the object in addition to the wanted holographic interference pattern. In figure 5 the relatively coarse structure is the recorded speckle pattern, which appears superimposed on the finer recorded holographic interference pattern.

2.2 The volume hologram

In section 2.1 it was assumed that the recorded representation of the interference pattern which constitutes the hologram is two-dimensional: that is, it exists entirely upon the plane defined by the sheet of recording medium (P in figure 3). A class of holograms exists, however, in which a volume of the interference pattern is recorded rather than a strict two-dimensional section. A practical photographic emulsion, for example, always has significant thickness compared with the wavelength of light and the resulting 'volume' hologram^{4e,19-21} will consist of surfaces of differing optical density (or differing refractive index if a phase hologram is produced) spaced throughout the thickness of the emulsion. Diffraction from such a hologram may be illustrated by considering the case discussed in section 2.1 where two beams of coherent light fall simultaneously onto the plate. In this case, however, the emulsion occupies the entire space shown in figure 3 and some additional notation is necessary (see figure 6) to describe the relevant relationships. The diagram relates to conditions within the emulsion of the photographic plate, where the refractive index (μ) is significantly greater than unity: thus $\lambda = \lambda_0 / \mu$, where λ_0 is the free-space wavelength. Furthermore, the angles between the beams differ from the free-space values because of

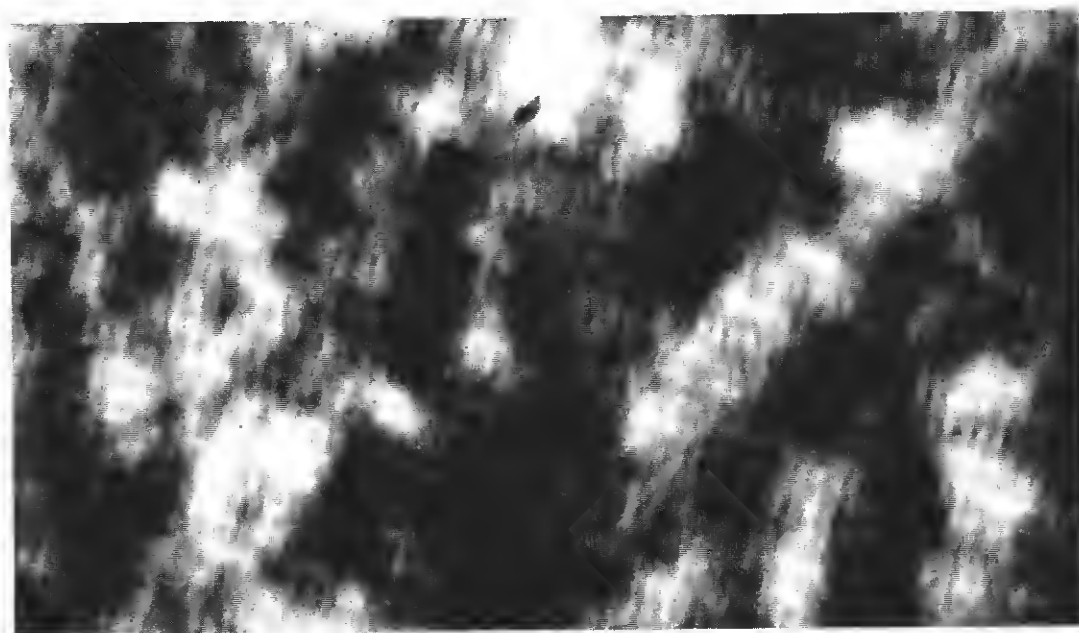


Fig. 5 Enlargement of a portion of a hologram

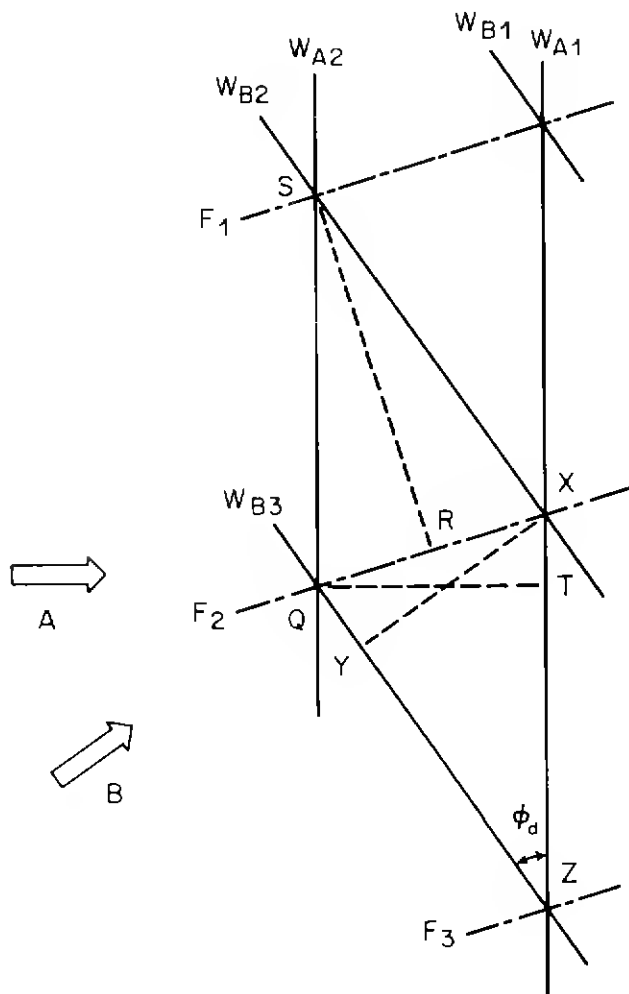


Fig. 6 Geometry of three-dimensional record of interference pattern

refraction at the emulsion and substrate surfaces. The following relationships then apply:-

- i) QT is perpendicular to SQ and defines the direction of beam A
- ii) XY is perpendicular to QZ and defines the direction of beam B
- iii) $XY = QT = \lambda$ where λ is the wavelength of the radiation (see above)
- iv) SR is perpendicular to QX and is of length d (the interference fringe spacing)
- v) $\angle X\hat{Z}Q = \phi_d$ (the angle between the directions of the two beams A and B)
- vi) The quadrilateral QXZS is a rhombus; hence $\angle S\hat{X}Q = \angle X\hat{Z}S = \angle S\hat{Q}X = \alpha$
- vii) $\angle X\hat{Q}T = \angle Q\hat{X}Y = \beta$, where β is the angle between the interference fringes and the direction of arrival of either beam A or beam B.

It can be seen that

$$\alpha = 90^\circ - \frac{\phi_d}{2} \quad (3)$$

and

$$\alpha = 90^\circ - \beta \quad (4) \quad \text{hence}$$

Also, since (equation 1)

$$XZ = \frac{\lambda}{\sin \phi_d}$$

the fringe spacing is given by

$$d = \frac{\lambda}{\sin \phi_d} \sin \alpha$$

$$= \frac{\lambda}{2 \sin \frac{\phi_d}{2}} \quad (5)$$

from equation 3. The processed hologram will therefore contain a number of parallel plane diffracting surfaces of spacing d as given in equation 5.

In considering the behaviour of the hologram during the image reconstruction process, use is made of a relationship derived (for X-ray diffraction from crystal lattices) by Bragg²². In figure 7 a set of parallel plane diffracting surfaces of spacing d is shown illuminated by a parallel beam of radiation of wavelength λ having a direction ψ relative to the diffracting planes. Bragg's Law states that the diffracted beam will emerge from the set of planes in such a direction as if specularly reflected by the planes (that is, so that the

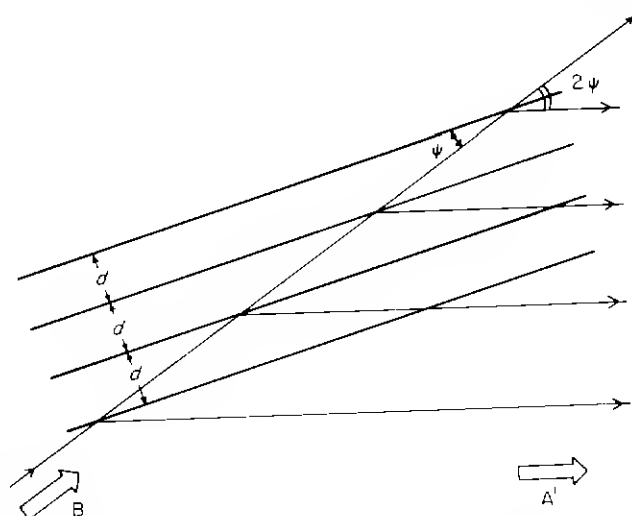


Fig. 7 Diffraction from three-dimensional recording

diffracted beam makes an angle 2ψ with the original beam). Furthermore, diffraction will only occur if the radiation components reflected from each plane are in phase: this leads to the relationship

$$2d \sin \psi = n\lambda$$

where n is an integer.

For first-order diffraction, $n = 1$ and

$$2d \sin \psi = \lambda.$$

Substituting for d from equation 5,

$$\psi = \phi_d/2 \quad (6)$$

or from equations 3 and 4

$$\psi = \beta \quad (7)$$

The diffracting planes shown in figure 7 occupy precisely the same positions within the emulsion of the photographic plate as the interference fringes from which they were derived during the formation of the hologram. Equation 7 thus implies that the angle of incidence of the reconstructing beam onto the hologram must be the same as that of one of the original beams when forming the hologram. (Figure 7 shows the case where the original beam B of figure 6 is used to reconstruct beam A, but the angular relationships are the same if beam A is used to reconstruct beam B.) The twin-wave beam will not be reconstructed although it can be obtained, by itself, by a different choice of reconstructing beam angle.

If, as discussed in section 2.1, several 'object' beams are incident at various angles onto the plate, in addition to the reference beam, separate sets of diffracting planes will be formed for each such object beam, and these beams will all appear during the reconstruction process, provided that the reconstruction beam approaches the hologram from the appropriate direction. As reconstruction will not take place if a volume is illuminated at other than the correct angle, it is possible to superimpose volume holograms on the same plate, using different reference beam angles, and recover only the wanted image by suitable choice of reconstruction-beam angle. The change in angle required to extinguish a reconstructed image depends on the number of diffracting planes present in the hologram: a detailed treatment of the subject is given by Smith^{4h}.

3 Characteristics of transmission and recording systems

In this section recording and transmission systems for sound and television are discussed in general terms, in order to provide background information for detailed discussion of holographic recording techniques.

3.1 General

When discussing recording systems, two classes of recording techniques may be distinguished:

- a) Analogue recording, in which the recorded information is related to the original information (e.g. sound pressure level or light flux from an elementary scene area) by a continuous mathematical function
- b) Digital^{23a} recording, in which the original information is sampled at suitable time intervals and the magnitude of each sample is recorded in numerical form.

In the case of the analogue recording of television signals, a further distinction may be made between 'picture' recording, in which the recorded information takes the form of recognisable images of the displayed picture²⁴, and 'waveform' recording in which the recorded information is an analogue of the electrical video signal^{25,26}. Sound analogue recording, whether electromechanical (disc), electromagnetic (tape) or electro-optic (film)²⁷⁻²⁹, is necessarily a 'waveform' process.

Since analogue recording techniques are very well established, any new system must show a worthwhile advantage over the more conventional methods in order to gain acceptance (see section 4.2). Digital recording techniques are less well established, although they have been under consideration for a number of years³⁰⁻³³. The process of 'digitizing' a signal is essentially the same for sound and television signals (see section 3.2) and no distinction can therefore be made between the basic digital recording techniques required for sound and television. In particular, a digital television recording does not provide a recognisable image of the displayed picture: however, in a system proposed by Hacking³⁴ for recording digital television signals, analogue information could in principle be re-introduced to facilitate editing, given a recording medium with suitable characteristics.

3.2 Summary of principles of digital transmission

In the process of 'analogue-to-digital conversion', the original signal must be sampled at least twice in the period of time occupied by one cycle of its highest permitted frequency^{23b}. A group of pulses (usually representing binary digits) can then be used to describe the magnitude of each such sample. The accuracy with which this sample magnitude can be described is determined by the number of binary digits in the group. If these pulses are transmitted in serial form (i.e. in succession through a single transmission path) the bandwidth of the resulting 'digital signal' will be proportional to the product of the sampling frequency and the number of digits (including parity digits³⁵) used to describe the sample magnitude and will thus be greater than that of the original analogue signal. It is convenient to define the required bandwidth in terms of the flow of binary digits or 'bits' through the transmission path. In the case of digital television signals^{30,36} for example, the flow is about 10^8 bits per second, while for a stereo sound signal³⁷ it is about 10^6 bits per second.

At the receiving terminal of a digital transmission system, each group of incoming binary digit pulses is used to generate a pulse of magnitude equal to the corresponding sample of the original analogue signal, within the limits of accuracy imposed by the number of digits in the group. The analogue output signal is obtained by low-pass filtering the resulting train of regenerated variable-amplitude pulses. The inaccuracies in regenerating the original sample magnitudes give rise to a noise-like component^{23c} added to the original analogue signal (the so-called 'quantising noise'). Provided that the signal-to-noise ratio of the digital signal at the input to the receiving terminal is sufficiently high to enable the presence or absence of a pulse to be recognised unambiguously, the quantising noise constitutes the only noise component of the regenerated signal. The effect of noise in the transmission channel is to cause errors in the recognition of incoming digital information. It has been found that, using appropriate error-concealment techniques, an error rate (i.e. the probability of erroneous recognition of any one digit) of 10^{-4} can be tolerated in a digital television system³⁸ and an error rate of 10^{-5} in a sound system³⁹, corresponding to ratios of peak signal to rms white noise in the

transmission channel of about 18dB and 19dB respectively^{23d}. These values may be compared with the signal-to-noise ratios of the regenerated analogue signals of 55dB for television⁴⁰ and 81dB for sound³⁹.

From the foregoing description it can be seen that a digital transmission channel requires a considerably wider bandwidth than the corresponding analogue channel, but can have a very much lower signal-to-noise ratio. Furthermore, the digital system lends itself to the transmission of signals over long distances, since a set of repeaters placed at intermediate points on the route can regenerate the digital information and in each case entirely eliminate the noise component of its incoming signal^{23e}. Distortion- and noise-free transmission (within the limits imposed by quantising noise) can therefore be obtained over such transmission paths.

3.3 General comments on digital recording

The facility, mentioned above, of digit-signal regeneration has led to interest in digital recording systems. By proper choice of the regeneration process, the reproduced signal can in principle be made free of all defects normally associated with analogue recording systems³¹. Unfortunately, recording systems are, by their nature, devices of limited bandwidth. The incoming information (whether analogue or digital) is retained in the recording medium as spatial variations of some property of the medium, and in all cases there is an upper limit to the spatial frequency which can be retained by the medium. This limit may be intrinsic to the medium itself, or may be due to limitations in the methods used to write the information onto the material or read it off again: in practice all such limitations are usually effective. In discussing digital recording techniques it is convenient to describe this limitation of the recording process in terms of the 'packing density' of the information on the recording medium, or in other words, the number of binary digits that can be recorded on (and recovered from) unit area of the recording medium. An estimate of the packing density that could in principle be achieved using a holographic digital recording system is made in section 5.

Digital recording systems may be used for both sound and television applications. Although the discussion which follows in sections 4-7 is directed towards the more difficult problem (in terms of rate of information flow) of recording digital television signals, digital techniques could also be of use for recording sound signals. In both cases, for example, copies of a digital recording held for archive purposes could be made at intervals, so as to offset the effects of long-term deterioration of the recording medium. This process is analogous to the insertion of repeaters in a long-distance transmission system (see section 3.2), and would enable the recording to be preserved without loss of quality.

4 The use of holographic techniques in recording systems

4.1 General

The property shown by a diffuse hologram of giving image reconstruction having considerable immunity to blemishes,

dust particles, etc. on the hologram itself has already been described (see section 1). This property can be illustrated further by considering the case of the holographic reproduction of a two-dimensional object situated in a plane parallel to, but separated from, that of the recording medium itself. The reconstructed image (I in figure 8) will

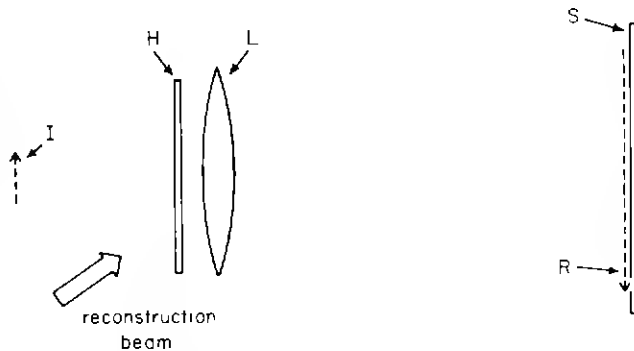


Fig. 8 Re-imaging of holographic image
 I — Reconstructed image
 H — Hologram
 L — Lens
 R — Final real image
 S — Screen

occupy a position relative to the hologram (H) identical to that of the object. If a lens (L) is placed close to the hologram, a final real image (R) of the reconstructed image may be formed on a screen (S), while the hologram itself will (if the lens is of sufficiently wide aperture) be well out of focus. Blemishes, etc. on the hologram will therefore not appear on the screen, and the image R will be free of defects from this cause. Such blemishes will, however, contribute to the overall flare or 'veiling glare' of the lens system. Furthermore, if only part of the hologram is illuminated by the reconstruction beam, the whole image will nevertheless be reproduced: the resolution of image detail will, however, be reduced since the effective area of the hologram will have been reduced to that of the illuminated portion.

4.2 Analogue recording

The reconstructed holographic image may be used as a basis for a system of 'picture' recording of television signals. Each television picture has first to be recorded on conventional motion-picture film and a hologram is then prepared from each film frame. Although at first sight this process appears unduly cumbersome, it confers the ability to use non-photographic materials as the holographic recording medium. Advantage may be then taken of the greater durability of these materials when making recordings for archival purposes: holograms could, for example, be prepared on metal surfaces. Assuming that a phase hologram is prepared (see section 1), it can be shown^{41,41} that if the hologram causes relatively small phase changes of the reconstructing-beam light, and if these phase changes are proportional to the holographic interference-pattern

intensity during the preparation of the hologram, then a distortion-free holographic image will be produced, i.e. the reproduction of the relative intensities of light from different parts of the holographic image will be correct. The first of these conditions can be achieved by suitable choice of exposure of the recording medium, while the second condition is ensured by utilising only a small part of the overall transfer characteristic (the relation between exposure and the resulting phase change of the reconstructing illumination) by arranging that the reference beam is of much greater intensity than the light from the object. Under these conditions the holographic image will remain distortion-free even if long-term changes take place in the transfer characteristic of the material. Because the object (film frame) is transparent, non-diffuse object illumination may be used and the speckle-pattern (see section 2.1) effects associated with the scatter of coherent light from diffuse objects will be absent. The advantages of the use of diffuse object illumination (see section 1) may be largely retained by illuminating the transparent object with a number of non-diffuse beams of light, inclined to each other at small angles. Each of these beams gives rise to a separate hologram on the recording medium^{41,42} and the geometry of the system is arranged so that these 'sub-holograms' do not overlap. The reconstructed images from each sub-hologram coincide to form the final reconstructed image, while the redundancy of information provided by the use of a number of sub-holograms gives protection against the appearance in the image of effects due to blemishes in the optical system or on the hologram itself. The required multiple-beam illumination may be obtained by the use of a mirror system, or by the use of a phase grating with one set of rulings ('one-dimensional grating') or two orthogonal sets of rulings ('two-dimensional grating') placed behind the object^{41,43,44}. Colour information may be carried by forming separate photographic images of each of the television colour-separation signals (or suitable coded versions of them) and making holographic recordings of each of these images. Alternatively, colour film may be used to form the intermediate object, and the colour information carried by illuminating the object and the recorded hologram with coherent light of (at least) three different wavelengths chosen in relation to the dye characteristics of the film. Arrangements would be required (for example, by using colour selective filters as in Collier and Pennington's method⁴⁵) to ensure that any one sub-hologram is formed by light of only one wavelength and is illuminated with light of the same wavelength during image reconstruction. Some electronic correction of the signals obtained by scanning the reconstructed images would probably be required to ensure reasonable colorimetric accuracy in the final reproduced picture.

The second method of recording colour information described above could also be used to make holographic recordings of conventional colour motion picture film for archive purposes: it is however likely that the directly-viewed reconstructed image would be subject to considerable colorimetric inaccuracy.

Apart from the advantage of extreme durability, which may, as discussed above, find an application in the preservation of pictorial information for archival purposes, the use

of holographic techniques does not seem to offer any benefits over other well-established methods of analogue television recording. Holography is essentially 'pictorial' in character, involving a physical object and a reconstructed image of that object, and implying the existence of a finite interval of time, however short, during which the object remains unchanged while information about it is impressed on to the recording medium. A truly continuous flow of information, as is present in an analogue 'waveform' signal, cannot therefore be recorded directly using holographic techniques.

4.3 Digital recording

In the basic digital holographic recording system a number of 'digit beams' of coherent light ($B_1, B_2 \dots$ in figure 9) are directed at different angles towards the recording medium^{40,46}. Each beam is associated with the digit having a particular degree of significance in a digit 'word': thus beam B_1 may correspond with the least significant digit in the word, beam B_2 with the next degree of significance, and so on. A reference beam (B_{ref}) is also incident onto the recording medium. All the beams are derived from the same coherent source by suitable light-splitting arrangements. Each digit beam can be either transmitted or obstructed by a light valve ($V_1, V_2 \dots$) controlled by the appropriate digit signal. The resulting hologram will consist of a number of superimposed fringe patterns, each pattern corresponding to one of the digit beams reaching the recording medium. For example, if each beam is accurately collimated, a set of grating-like patterns will be produced, each having a different combination of spatial frequency and orientation. In figure 9 the digit word 1101 is shown being recorded, beam B_2 being obstructed and the other beams transmitted by their respective light valves.

When the processed hologram is illuminated by a reconstruction beam (B_{rec} in figure 10), diffraction occurs at each of the fringe patterns and the digit beams incident during the recording are reconstructed ($B'_1, B'_2 \dots$). The twin-wave beams $B''_1, B''_2 \dots$ may also be produced^{4a} and the residual reconstruction beam B_{rec} will also be present. Photodetectors $P_1, P_2 \dots$ sense the presence or absence of each reconstructed digit beam.

During the formation of the hologram mutual interference will also take place between the digit beams and the patterns thus recorded on the hologram will give rise to spurious 'intermodulation' diffracted beams on reconstruction (not shown in figure 10). Their intensity may in some cases be enhanced by geometrical symmetry in the arrangement of digit beams (for example, if the angles between adjacent digit beams are all equal, the intermodulation components due to interference between adjacent beams will all be superimposed to form one relatively intense spurious component). This effect may be reduced by providing⁴⁷ a random phase relationship between the various digit beams arriving at the recording medium, through appropriate adjustments of the optical path lengths. In any case, the geometry of the system must be arranged so that no spurious beam of significant intensity can fall onto a photodetector.

Instead of using light valves to control the digit beams

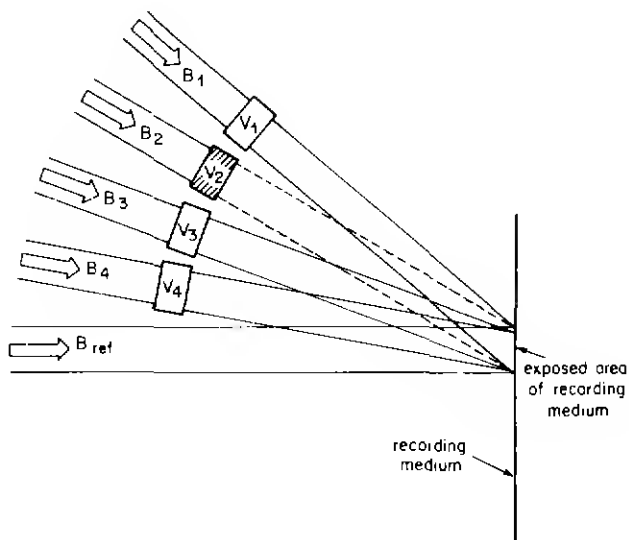


Fig. 9 Making a digital holographic recording. basic principles

B_1, B_2, \dots — Digit beams
 B_{ref} — Reference beam
 V_1, V_2, \dots — Light valves

during the recording process, a fringe-spoiling technique may be used in which the position of the relevant interference fringes relative to the recording medium is changed during the exposure of the medium^{34,46}. A blurred recording of the interference fringes is thus obtained resulting in a reduced modulation depth of the recorded pattern. The extent of this reduction depends on the amount and nature of the movement of the fringe pattern⁴¹ but can be arranged to eliminate the recording of the pattern completely. The corresponding digit beam will therefore not be reconstructed. Another method of digit-beam control consists of arranging that the reference beam is plane polarised, and that the polarisation of each digit beam can be either parallel to or perpendicular to that of the reference beam, so that interference between the two beams does or does not take place. Both these methods have the advantage that the mean exposure of the recording medium does not depend on the recorded digital word: this is an advantage since work by Hacking has shown³⁴ that only a rather restricted latitude of exposure of the medium ($\pm \frac{1}{2}$ stop) can be permitted for satisfactory recording of the interference fringes.

In practice, the recording medium itself must move between successive exposures, and arrangements must be made to ensure that this movement, if continuous, does not significantly reduce the modulation depth of all the recorded interference patterns. One such arrangement is to use a very short exposure time, e.g. by using a Q-switched laser^{9,48}; such an approach, however, would make the fringe-spoiling technique discussed above very difficult to implement. Alternatively, the movement of the medium can be made parallel to the individual lines in the fringe patterns, although this imposes restrictions on the directions of arrival of the digit beams.

5 The maximum packing density of a holographic recording system

In a holographic recording system, a succession of holo-

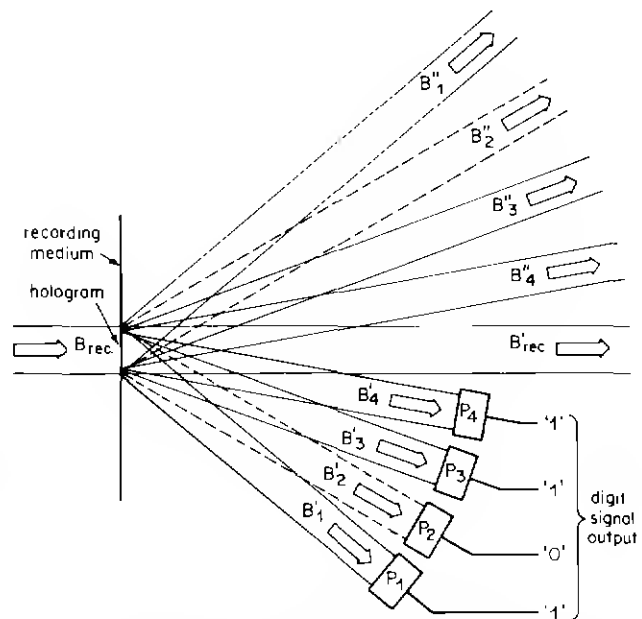


Fig. 10 Replaying a digital holographic recording. basic principles

B_{rec} — Reconstruction beam
 B'_1, B'_2, \dots — Reconstructed digit beams
 B''_1, B''_2, \dots — Reconstructed twin wave beams
 P_1, P_2, \dots — Photocells

grams will be recorded side-by-side on the recording medium. The calculation of the packing density of the recording system (see section 3.3) therefore resolves into obtaining the product of the number of binary digits in each hologram and the number of holograms on unit area of the recording medium.

The information capacity of a single hologram clearly depends on the size of the hologram and the resolution of the recording material (i.e. the highest spatial frequency that the material can accommodate). A number of authors^{4,49,50} have discussed this subject: the following treatment is based on the method used by Smith^{4k}.

Consider a square, transparent surface hologram of side H , uniformly illuminated by a collimated reconstruction beam normal to the surface of the hologram. Let λ be the wavelength of the illuminating radiation. Assume that the recorded interference patterns consist of sets of parallel straight lines, so that the reconstructed beams would also be collimated if H was very large.

In practice H is not large, and diffraction occurs because the hologram acts as finite aperture. Each reconstructed beam will therefore diverge: at distances large compared with H the intensity profile of each beam will contain maxima and minima. Using the Rayleigh resolution criterion, two adjacent beams (B_1 and B_2 in figure 11) are considered distinguishable from each other when the direction of maximum intensity of one beam coincides with the direction of the first minimum of intensity of the other beam. Under these conditions the angle ϕ_b between the beams is given^{14,51} by

$$\sin \phi_b = \frac{\lambda}{H} \quad (8)$$

(Strictly speaking this equation only applies to sets of beams in planes parallel to the hologram edges and in directions nearly normal to the hologram surface: however, for the present order-of-magnitude estimate it will be taken as

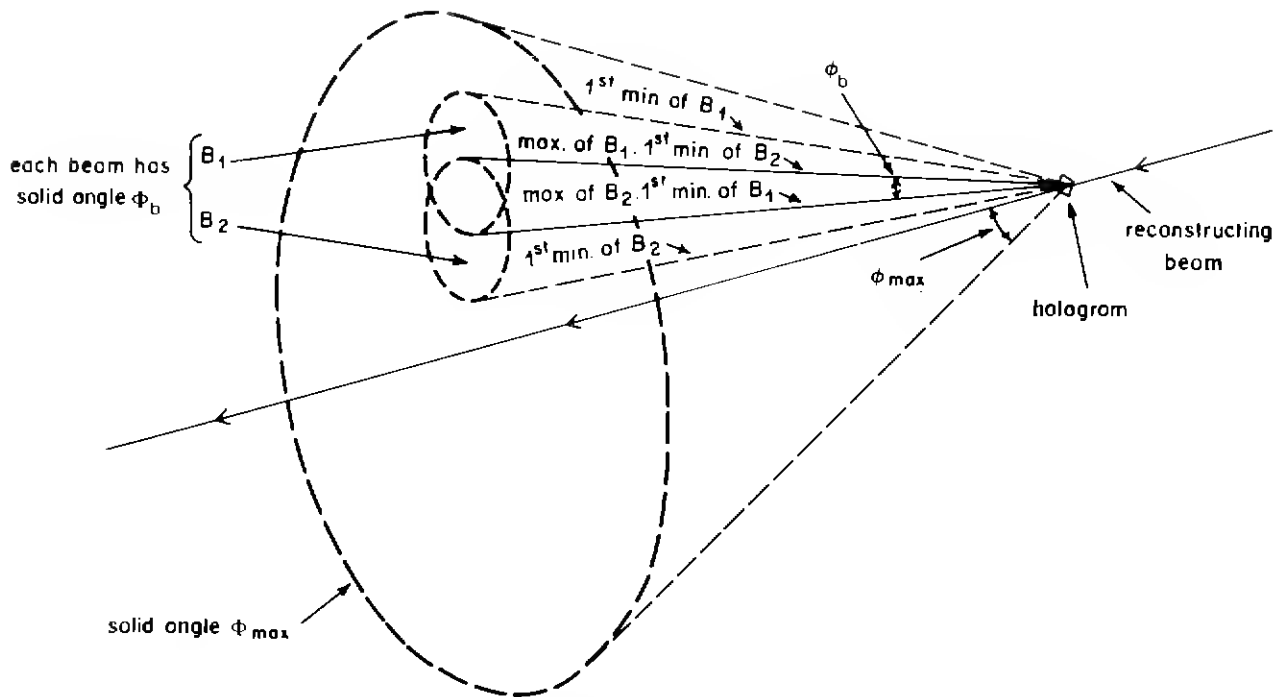


Fig. 11 Cones of radiation from a small hologram with limited resolution

applying to all beams emerging from the hologram.) The solid angle Φ_b occupied by one beam is given by

$$\Phi_b = 2\pi (1 - \cos \phi_b).$$

If the resolution of the recording medium is restricted to a spatial frequency of n_{max} , it can be shown (equation 2) that first-order diffracted beams cannot emerge from the hologram at an angle to the normal greater than ϕ_{max} where

$$\sin \phi_{max} = \lambda n_{max}. \quad (9)$$

The solid angle Φ_{max} which may be occupied by reconstructed beams is therefore given by

$$\Phi_{max} = 2\pi (1 - \cos \phi_{max}).$$

If adjacent beams are considered as just touching, the number of beams (N) that can occupy this space is given by

$$N = \frac{\Phi_{max}}{\Phi_b}.$$

Hence, applying the Rayleigh resolution criterion so that the beams overlap as shown in figure 11, the greatest number of resolvable beams (N_{max}) that can occupy this space becomes

$$\begin{aligned} N_{max} &= \frac{4\Phi_{max}}{\Phi_b} \\ &= \frac{4(1 - \cos \phi_{max})}{1 - \cos \phi_b}. \end{aligned} \quad (10)$$

In practice, the hologram dimension H is likely to be much greater than the wavelength λ of the illuminating radiation:

hence, from equation 8, $\sin \phi_b$ (and therefore ϕ_b) is very small, and this equation may be written

$$\sin \phi_b = \phi_b = \frac{\lambda}{H}$$

furthermore, because ϕ_b is small,

$$1 - \cos \phi_b = \frac{(\phi_b)^2}{2}.$$

Hence equation 10 becomes

$$N_{max} = \frac{8H^2 (1 - \cos \phi_{max})}{\lambda^2}.$$

Since a surface hologram is under consideration, each interference pattern recorded on it will give rise to two reconstructed beams, corresponding to the wanted and twin-wave holographic images^{4a}. Hence the number of different interference patterns that can be recognised (and therefore the number of binary digits, B_{max} , that can be recorded on the hologram) is given by

$$B_{max} = \frac{N_{max}}{2} = \frac{4H^2 (1 - \cos \phi_{max})}{\lambda^2}. \quad (11)$$

Since $\sin^2 x + \cos^2 x = 1$, equation 9 may be written

$$\cos \phi_{max} = [1 - (\lambda n_{max})^2]^{\frac{1}{2}}.$$

Hence equation 11 becomes, in terms of the original maximum spatial frequency n_{max} that can be laid down on the recording medium

$$B_{max} = \frac{4H^2}{\lambda^2} \left\{ 1 - [1 - (\lambda n_{max})^2]^{\frac{1}{2}} \right\} \quad (12)$$

An interesting feature of equation 12 is that it shows the number of binary digits that can be carried in one hologram to be proportional to H^2 or, in other words, to the area of the hologram. The theoretical packing density of the holographic recording system is thus independent of the size of the individual holograms, provided that no recording-medium surface is left unexposed between these holograms. For helium-neon radiation ($\lambda = 632.8\text{nm}$) equation 12 leads to the result that for a recording medium limited in resolution to a spatial frequency of 1000 cycles/mm the greatest possible packing density is of the order of 2×10^9 bits/cm².

Work by Langdon⁵⁰ has shown that an acceptably low error rate in detecting the presence or absence of a particular digit beam can only be obtained if the separation between adjacent beams is greater than that predicted by the Rayleigh criterion. This is because 'spatial noise' or random fluctuation in intensity from point to point in the reconstructed image, which can occur for a variety of reasons including speckle-pattern effects, deficiencies in the light valves used to make the hologram and vignetting produced by the optical components in the recording and replay equipment, causes individual beams to differ considerably in intensity. This gives rise to difficulties in choosing a threshold intensity level. Langdon found that the area of hologram required to give adequate separation of the digit beams was four times the value predicted by the Rayleigh criterion. This suggests a practical upper limit on the packing density which is lower than the theoretical maximum by a factor of four: in these terms the practical limit for a recording medium with resolution limited to 1000 cycles/mm would be 5×10^7 bits/cm². In fact, Langdon did not exploit the full range of resolution capabilities of his storage medium, but by the use of an off-axis reference beam used a relatively narrow range of spatial frequencies. In consequence the packing density that was achieved (10^4 digits in a hologram 0.6mm in diameter) was only about one fifteenth of the above-described practical upper limit. This further limitation is purely instrumental in character and in principle could be overcome, although it is possible that spatial noise could pose an increasingly serious problem as the number of reconstructed digit beams was increased. To allow for this effect a further reduction by a factor of five in the practical upper limit on packing density is suggested, giving a final practical upper limit of 10^7 bits/cm².

The discussion so far has been concerned with surface holograms. If the recording medium is suitable for preparing volume holograms, Bragg-angle effects⁴¹ can be used for two purposes. In the first place reconstruction of the unwanted twin-wave digit beams is suppressed, and the storage capacity of the hologram is therefore doubled. In addition, a number of holograms can be superimposed during recording, using differing reference-beam angles of incidence for each individual hologram: image reconstruction will only occur if the reconstruction- and reference-beam angles of incidence correspond. The necessary degree of such correspondence depends (among other factors) on the thickness of the recorded hologram. For a thickness of 100 wavelengths differences in reconstruction-beam angles of incidence of the order of four degrees^{4b} will enable the individual holograms to be distinguished from each other.

This suggests that the limit to the number of individual holograms that can be superimposed will be set by the transfer characteristic of the recording medium rather than by restrictions in the choice of reference-beam angle. The superposition of up to seven analogue holograms has been described⁵²: it seems reasonable to assume that at least ten digital holograms could be superimposed. This would increase the theoretical packing density by a factor of 20, to 4×10^9 bits/cm², while the practical upper limit, as discussed above, would rise to 2×10^8 bits/cm² (although reservations on the achievement of this packing density must be made because of the possibility that the effect of spatial noise would be still further enhanced).

The use of volume holograms becomes even more advantageous if very thick recording media are used. Van Heerden has estimated⁵³ that a total of 10^{13} bits/cm³ may be stored in a suitable crystalline material. It does not appear that such a high packing density has yet been achieved in practice, although a system giving a potential storage capacity of 10^{12} bits/cm³ has been announced⁵⁴.

6 Practical considerations

6.1 Optical arrangements

A number of arrangements may be used for generating the individual digit beams (figure 9). One method is to use a relatively coarse diffraction grating, giving a uniform distribution of energy through many orders of diffraction: an array of small parallel cylindrical lenses embossed on a transparent film has proved suitable for this purpose⁴⁶. If the number of digits per hologram is very large, a more convenient arrangement is to provide a digit plate (P in figure 12) consisting of an array of holes in an otherwise opaque medium⁵⁰. This plate is rear-illuminated with a beam of coherent light (B_{illum}) and is placed in the focal plane of the object lens (L_0). Some method is provided for causing the light passing through each hole to diverge: for example, a diffuser may be placed behind the plate, or the plate may be illuminated with a number of non-diffuse beams inclined to each other as discussed in section 3.2. Alternatively, enough divergence of the light may be obtained by diffraction at the digit-plate holes. Each hole in the digit-plate acts as an individual source and gives rise to a quasi-collimated beam* (B_1, \dots, B_n) after passing through the object lens, the angle made by the beam to the axis of the optical system being determined by the position of the hole in the digit plate. These beams converge at the focus (F_0) of the object lens, where the holographic recording medium (R) is placed. A reference beam (B_{ref}) of greater intensity than the digit beams also falls on the holographic recording medium at this point. (Although shown as emanating from a separate source in figure 12, it could be generated from another hole in the digit plate with suitable adjustment of relative intensities.)

The system described by Langdon⁵⁰ involved the simulation of a read-only memory, the digit plate consisting of an array of transparent square areas on an opaque background, some of the transparent areas being blacked-out to simulate logical 'noughts'. In a fully-developed, read-only *Strictly speaking the beam is collimated only when the digit plate hole is sufficiently small to be regarded as a point source.

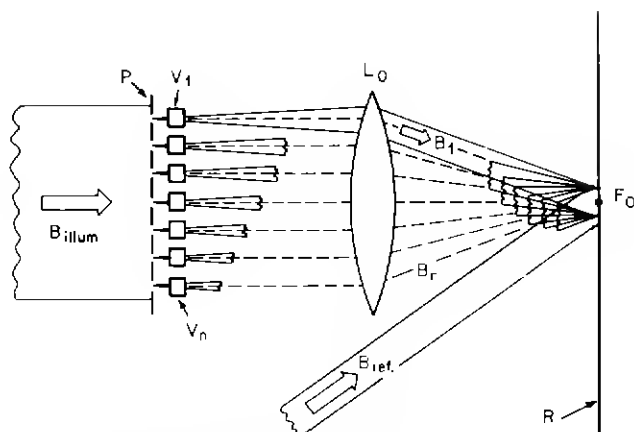


Fig. 12 Making a digital holographic recording with a large number of digit beams

P — Digit plate
 B_{illum} — Illuminating beam
 B_{ref} — Reference beam
 B_1, \dots, B_n — Digit beams
 V_1, \dots, V_n — Light valves
 L_0 — Object lens
 R — Recording medium
 F_0 — Focus of object lens
 (For clarity only digit beam B_1 is shown in full)

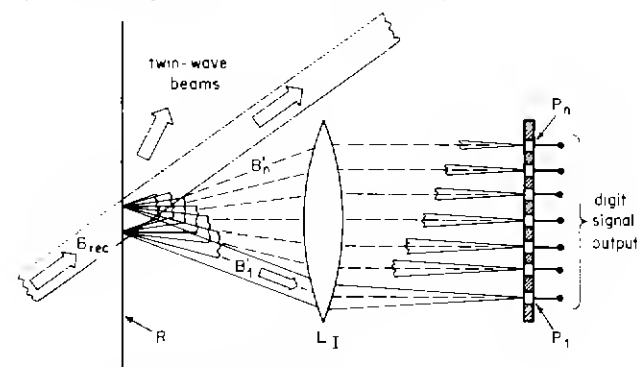


Fig. 13 Replaying a digital holographic recording using a large number of digit beams

B_{rec} — Reconstruction beam
 B'_1, \dots, B'_n — Reconstructed digit beams
 R — Recording medium
 L_I — Imaging lens
 P_1, \dots, P_n — Photocells
 (For clarity only reconstructed digit beam B'_1 is shown in full and the twin-wave beams — cf. Fig. 3 — are not shown in detail)

memory a number of digit plates would be used (one per recorded hologram), each with the appropriate pattern of digital 'ones' and 'noughts' drawn on it. For a real-time recording system, all the holes in the digit plate would be present and an array of light valves (V_1, \dots, V_n) adjacent to the plate would control each digit beam.

The photocells (P_1, \dots, P_n in figure 13) used to recover the digital signals from the reconstructed digit beams (B'_1, \dots, B'_n) during playback of the recording are most conveniently constructed in an integrated array if the number of digits per hologram is large. A lens system (L_I) is therefore required to image the reconstructed digit beams onto this array. With a small number of digit beams it would be feasible to use separate photocells placed directly in the paths of the reconstructed digit beams. The reconstruction beam (B_{rec}) is incident onto the holographic recording medium (R) at the same angle as the reference beam used in the recording process. In fact, the same equipment could be used for both recording and playback: during playback the beam B_{illum} (figure 12) would be absent, while during recording the photocells (P_1, \dots, P_n in figure 13) would

provide a monitoring facility if the recording medium was partially transparent.

The apertures of the lens systems used during recording and playback impose restrictions on the greatest angle of incidence of the digit beam which can be used during recording and the greatest angle of emergence of the reconstructed digit beam which can be accepted during playback. It may be necessary to use a number of suitably disposed optical systems, with the axis of each system inclined to the axis of the complete equipment (i.e. the normal through the surface of the recording medium), in order to make full use of the resolution capabilities of the recording medium.

6.2 Recording and playback

The speed of the recording medium, if used in the form of a continuous strip or tape, is determined by the information flow rate, the packing density and the width of the track laid down on the recording medium. An order-of-magnitude estimate of this speed may be obtained by assuming that Langdon's result⁵⁰ of holding 10^6 binary digits in a hologram 0.6mm in diameter can be achieved. As the information flow rate of a digital television signal is 10^8 bits/second, 10^6 holograms per second will be required and the tape speed will therefore be 6m/sec (240 in/sec) if the holograms are arranged just to touch each other and are laid down in a single line. This tape speed is excessive and moreover the tape width (say 1mm or $\frac{1}{2}$ inch) required for the economical use of the recording medium would make the medium very intractable.

A more practicable recording arrangement would be obtained if the individual holograms were laid down in a series of transverse tracks across a strip of recording medium of more manageable dimensions. For example, a one-inch (2.5 cm) wide tape could be used; forty holograms could then be accommodated in each 0.6mm-wide track giving a tape velocity of 15cm/sec (6in/sec). This technique would involve a method of scanning the recording or reading area transversely across the tape such that the full width of the tape was traversed in four milliseconds (or in other words 250 traverses per second occurred). Methods of achieving such scanning action have been described by Childs⁵⁵; of particular interest in the present context is the use of a binary sequence of birefringent elements^{56,57}. During the recording process the exposure time would be short enough to 'freeze' the tape motion so that the recorded interference patterns were not blurred. (For a tape speed of 15cm/sec the tape moves about one tenth of a wavelength in 400n sec.) On replay, the lens L_I (figure 13) would have a diameter which was larger than the tape width, so that all reconstructed digit beams were collected. It is possible to arrange the optical path so that light from each reconstructed digit beam always falls onto the appropriate photocell, irrespective of the position across the tape of the illuminated hologram.

If a packing density of 10^7 bits/cm² could be achieved (see section 5) the tape velocity could be reduced by a factor of three relative to the value discussed above. A further reduction in tape velocity could also be made if the recording medium was capable of accepting volume holograms, and multiple exposure of the medium therefore became possible: the velocity reduction factor would be equal to the number of holograms that could be superimposed. Combining these

two factors and assuming that ten holograms could be superimposed, the tape velocity required for the recording of a digital television signal onto one-inch (2.5cm) wide tape, as discussed above, would become 0.5cm/sec ($\frac{1}{4}$ in/sec).

If the potentialities of the use of crystalline materials could be realised (see section 5) an even more compact form of storage of digital signals would become available. The system⁵⁴ capable of storing 10^{12} bits/cm³, for example, could store 10^6 seconds (about 24 hours) of digital television programme, or 10^8 seconds (about 11½ days) of stereo sound signal, per cubic centimetre of recording material.

6.3 Hologram size

Apart from the high packing density that can be achieved in a holographic digital recording system, two other advantages of this system are important. In the first place, because the holographic pattern corresponding to each binary digit covers the entire surface area of an individual hologram, defects due to dust, blemishes etc. on the recorded hologram will not cause 'drop-outs' (i.e. the complete loss of one or more digit signals) to occur unless the defects are so large as to obscure a major part of the hologram. The effect of relatively small defects will instead show as a reduction in signal-to-noise ratio (i.e. a reduction in the contrast between the digit beam and the background flare illumination) which will have no effect on the recovery of digital information until a certain threshold is reached. This advantage clearly becomes increasingly significant as the size of the individual holograms is increased. In the second place, a reduction in signal-to-noise ratio can also occur if the reconstruction beam falls onto a hologram adjacent to the one intended. Again this will have no effect on the reproduced digital signal until a definite tolerance limit is exceeded. The reduction in signal-to-noise ratio will in this case depend on the relative areas of the 'intended' and 'unintended' holograms illuminated by the reconstruction beam. As the size of the individual holograms is increased, the absolute dimensional accuracy with which the reconstruction beam must locate a hologram is therefore reduced, making for easier mechanical and optical design of the equipment. Both these considerations lead to the conclusion that the size of each individual hologram should be as large as possible, so as to give the greatest benefit in terms of immunity from recording medium defects and relaxation of dimensional tolerances.

Although, on theoretical grounds (see section 5), the packing density of the recording system is independent of the individual hologram size, the use of large individual holograms in a recording system of given packing density implies the use of a correspondingly large number of digit beams. Thus as hologram size is increased, basic optical and mechanical tolerances (e.g. recording medium blemishes and location tolerances on replay) become less stringent, but only at the expense of increased electro-optic complexity in the generation and recognition of individual digit beams. The practical compromise between these two opposing factors is discussed in section 6.4 below.

6.4 Present-day feasibility of implementing a holographic digital recording system

It is unfortunately true that at present the use of a very large number of digit beams is not feasible. The principal limita-

tion is the absence of a suitable method of digit-beam control. For the simultaneous control of a large number (up to 10^4) of digit beams some form of integrated array of light valves is required. Work on the development of such arrays using liquid crystal cells⁵⁸ or a moving membrane⁵⁹ has been reported, but no device is at present available having the required bandwidth. It is possible that 'magnetic bubble' domains⁶⁰ might prove useful as the basis of such an array⁶¹, since the domains may readily be created, moved and annihilated by currents in suitably-disposed conductors. They may be used for digit beam control by virtue of the differential rotation of the polarisation of radiation passing through them relative to the radiation passing through the bulk of the magnetic material.

The use of a large number of digit beams also implies that the photo-cells used during replay should also be in the form of an integrated array, so that the image formed by the imaging lens (L_1 in figure 13 is of reasonable size). However, since such arrays find applications in solid-state television cameras, as well as in holographic read-only computer memories⁵⁰, much effort has been directed to their development, and in fact arrays containing 2×10^5 photo-transistors (400 rows and 500 columns) have been constructed⁶². It is not so certain, however, whether the addressing circuits associated with these arrays would at present be capable of handling the digital television information flow rate of 10^8 bits per second.

From the foregoing discussion it appears unlikely that a system exploiting in full the advantages of holographic digital storage can be developed in the immediate future. It would nevertheless be possible to consider a system (such as the one described by Hacking³⁴) using a practicable number of digit beams and at the same time using an individual hologram size such as to give some degree of protection against dust particles and misdirection of the reconstructing beam. Although the packing density of such a system would be lower (probably by several orders of magnitude) than the greatest value offered by a holographic system, it might still possess advantages over other currently-available methods of digital recording (see section 7) and in any case, its construction would provide valuable insight into the problems associated with holographic recording.

7 Comparison of holographic and other proposed digital recording systems

Digital recording systems are widely used for computer data storage, but no fully-developed system capable of handling the digital television information flow rate of 10^8 bits per second appears to exist at the present time. Several systems do, however, exist which might be capable of development to the point at which this information flow rate could be handled. These have been discussed by Lunn and Moffatt³⁰, and more recently by Jones³³. One interesting system uses a focused high-power laser beam to burn clear areas corresponding to individual binary digits in a specially developed opaque recording medium. The system as described uses holes of 1.5µm diameter separated by 3µm and laid down in rows 4µm apart (using a helical-scan system) and thus has a packing density of 8.3×10^6 bits/cm²: the information flow rate is stated to be 2×10^7 bits per second. It is claimed that

the system has a potential packing density of 10^8 bits/cm² but that additional development work would be required for the recording of digital television signals. If this potential performance was achieved the packing density of the system would be about an order of magnitude higher than that estimated to be practicable (see section 5) in a system using surface holograms: it would, however, have the disadvantage of requiring very accurate relocation of the individual binary digits on the recording medium during replay and would seemingly be very susceptible to drop-out errors due to dust, etc. on the recording medium (although an error rate of 1 in 10^9 over the whole recording, storage and read-out process is claimed). In the field of magnetic tape systems, equipment has been developed³², based on longitudinal multi-track techniques, which achieves an information flow rate of 71×10^6 bits per second, and yields a packing density of 10^5 bits/cm². It is hoped that this technique can be extended to handle the full television information flow rate of 10^8 bits per second.

Hacking's proposal³⁴ for a holographic recording system based on present-day technology uses small holograms recorded on conventional high-resolution film, and achieves a packing density of 10^6 bits/cm² (about an order of magnitude greater than that of the magnetic tape system described above). A considerable degree of accuracy in relocating the individual holograms is still required by this system, although the use of the holographic technique permits about an order of magnitude less accuracy in both directions (i.e. between individual holograms and between rows of holograms) compared with the 'individual' binary-digit system.

8 Recapitulation and conclusions

A hologram is a replica of the interference pattern formed in the common volume occupied by two or more coherent beams of radiation. One of these (the 'reference' beam) has wavefronts of well-defined form (e.g. plane or spherical with known centre of curvature) while the others come from an object illuminated by the same source. A 'surface' hologram is two-dimensional and replicates a section of the interference pattern, while a 'volume' hologram, as its name implies, is three-dimensional and replicates the interference pattern in depth. When a hologram is illuminated by a 'reconstructing' beam identical to the reference beam, the spatial amplitude and phase distribution of one component of the diffracted radiation is identical to that of the original radiation emanating from the object. A virtual image of the object is therefore formed, occupying the same position (relative to the hologram) as did the original object, and having an appearance (including its three-dimensional nature) identical to that of the original object. Other diffracted components of radiation may also be present, giving rise to spurious images.

If the object or the illuminating radiation is diffuse, radiation from any one point on the object will propagate through the complete volume in which the interference pattern is formed. Information about this object point will thus be distributed over the whole area (or volume) of the resulting 'diffuse' hologram. During the image recon-

struction process, a diffracted component of radiation corresponding to this original object point will therefore emerge from every part of the diffuse hologram: thus the presence of dust particles, blemishes on the hologram, etc. will not cause the entire loss of an item of information held by the hologram, as would be the case in a system (such as conventional photography) where such items are localised in specific areas of the record.

Diffuse objects illuminated by coherent light appear to be overlaid by a random-noise pattern (the 'speckle pattern') which is stationary if both object and observer are stationary. Speckle pattern also appears in the reconstructed image of such an object. This pattern is due to interference between components of radiation scattered from different parts of the object.

Holograms can be formed on a number of light-sensitive materials, some of which offer advantages over the use of photographic film. In particular, materials which introduce point-to-point changes in the phase of the incident radiation during image reconstruction, instead of corresponding amplitude changes, may be used. This technique widens the choice of materials which can be used as the recording medium, and also makes possible greater 'diffraction efficiency' (i.e. the proportion of the radiant energy incident onto the hologram during image reconstruction that is utilised in the wanted image-forming diffracted component) than is possible when the hologram introduces amplitude changes of the reconstructing radiation.

An analogue television recording system could be developed using holographic techniques. Such a system would not, however, offer significant advantages over present-day analogue systems, except perhaps in the use of recording materials of greater durability.

Since the information flow rate in digital television signals is very high (10^8 bits per second), it is essential to achieve as high a packing density as possible to keep the recording-medium transport velocity (e.g. tape speed) and the bulk of recording-medium used within practicable limits. Individual digits, if recorded in discrete form, will each occupy a very small area of the recording medium, and the reproduced signal will be very susceptible to errors arising from masking by dust particles and defects on the recording medium. A holographic system of recording would enable the digital information to be spread over an area of recording medium which is large compared with dust particles, etc. and thus would greatly reduce the errors in reproduction due to this cause. The use of a holographic technique also reduces the positioning accuracy required between the recorded information and the reproducing equipment. These advantages are however achieved only at the expense of increased optical and electronic complexity. A very high packing density can, in principle, be achieved using a holographic digital recording system, thus leading to low values of recording-medium velocity. Although an optimised holographic recording system cannot at present be implemented because some optical components are not available in a suitable form, it would be practicable to develop a prototype system using existing components. Valuable experience would thus be gained in this field, and a system could emerge which, although not representing the best that could

ultimately be achieved by the use of holographic techniques, might still be better than, say, magnetic-tape systems in terms of packing density and error rate in the reproduced signal. In the long term, it seems possible that a holographic digital recording system would enable extremely high packing densities to be obtained.

It does not appear feasible to develop a holographic television transmission system of broadcast quality.

9 References

Because the published material on holography is now very extensive, the references to this subject given below should be regarded as introductions to the topics concerned. Reference 4 is recommended as a concise treatment on a wide range of holographic techniques. Reference 19 treats the subject in greater detail and from a more general mathematical point of view; it also includes reprints of references 1, 2 and 3. Both books contain extensive lists of references, as do most of the cited papers. The RCA Review of March 1972 (see references 20 and 21) gives a comprehensive treatment of holographic recording materials and also contains discussions of general holographic theory. Recent papers are listed in section 0.648 and 2.460 (08.27 before January 1973) of 'Current Papers in Physics' and in section 2.970 (17.80 before January 1973) of 'Current Papers in Electrical and Electronics Engineering', both published by the Institution of Electrical Engineers.

- 1 Gabor, D. A new microscopic principle. *Nature*, 1948, Vol. 161, 4098, pp. 777-778.
- 2 Gabor, D. Microscopy by reconstructed wave-fronts. *Proc. Roy. Soc.*, 1949, Vol. A197, 1051, pp. 454-487.
- 3 Gabor, D. Microscopy by reconstructed wave-fronts: II. *Proc. Phys. Soc. Section B*, 1951, Vol. 64 Part 6, 378B, pp. 449-469.
- 4 Smith, H.M. *Principles of holography*. New York, Wiley-Interscience, 1969.

4a. pp. 3-6	4g. Chap. 4
4b. p. 25	4h. pp. 71-73
4c. pp. 48-53	4i. p. 53
4d. p. 49	4j. Section 6.2
4e. pp. 13-14	4k. pp. 220-222
4f. pp. 16-21	4l. pp. 63-67
- 5 Bragg, W.L. The X-ray microscope. *Nature*, 1942, Vol. 149, 3782, pp. 470-471.
- 6 Leith, E.N. and Upatnieks, J. Reconstructed wavefronts and communication theory. *J. Opt. Soc. Am.*, 1962, Vol. 52, 10, pp. 1123-1130.
- 7 Leith, E.N. and Upatnieks, J. Wave front reconstruction with diffused illumination and three-dimensional objects. *J. Opt. Soc. Am.*, 1964, Vol. 54, 11, pp. 1295-1301.
- 8 Sinclair, D.C. and Bell, W.E. *Gas laser technology*. New York, Holt, Rinehart and Winston, Inc., 1969.
- 9 Brown, R. *Lasers; a survey of their performance and applications*. London, Business Books Ltd., 1969.
- 10 Leith, E.N., Upatnieks, J., and Haines, K.A. Microscopy by wavefront reconstruction. *J. Opt. Soc. Am.*, 1965, Vol. 55, 8, pp. 981-986.
- 11 Meier, R.W. Cardinal points and the novel imaging properties of a holographic system. *J. Opt. Soc. Am.*, 1966, Vol. 56, 2, pp. 219-223.
- 12 Kogelnik, H. Reconstructing response and efficiency of hologram gratings. *Proc. Symposium on Modern Optics*, New York, March 22-24, 1967. (Microwave Research Institute Symposia Series, Vol. XVII), Polytechnic Press, 1967, pp. 605-617.
- 13 Kogelnik, H. Hologram efficiency and response. *Microwaves*, 1967, Vol. 6, 11, pp. 68-73.
- 14 Ditchburn, R.W. *Light*. London, Blackie, 1963.
- 15 Burckhardt, C.B. Efficiency of a dielectric grating. *J. Opt. Soc. Am.*, 1967, Vol. 57, 5, pp. 601-603.
- 16 Upatnieks, J. and Leonard C. Efficiency and image contrast of dielectric holograms. *J. Opt. Soc. Am.*, 1970, Vol. 60, 3, pp. 297-305.
- 17 Lin, L.H. Method of characterizing hologram-recording materials. *J. Opt. Soc. Am.*, 1971, Vol. 61, 2, pp. 203-208.
- 18 Taylor, E.W. An appraisal of the use of holographic techniques in television systems of broadcast quality. BBC Research Department Report No. 1973/39.
- 19 Stroke, G.W. *An introduction to coherent optics and holography*. London, Academic Press Inc., 1969.
- 20 Ramberg, E.G. Holographic information storage. *RCA Rev.*, 1972, Vol. 33, 1, pp. 5-53.
- 21 Meyerhofer, D. Phase holograms in dichromated gelatin. *RCA Rev.*, 1972, Vol. 33, 1, pp. 110-130.
- 22 Bragg, W.L. *The crystalline state. Vol. 1 — a general survey*. London, Bell, 1933.
- 23 Bell Telephone Laboratories. *Transmission systems for communications*, Winston-Salem, Bell Telephone Laboratories, Inc., 1964.

23a. Chapter 24
23b. pp. 597-598
23c. pp. 612-613
23d. p. 609
23e. Chapter 26
- 24 Palmer, A.B. The technical problems of television film recording. *J. Soc. Motion Pic. Tele. Engrs.* 1965, Vol. 74, 12, pp. 1069-1084.
- 25 Ginsburg, C.P. Comprehensive description of the Ampex video tape recorder. *J. Soc. Motion Pic. Tele. Engrs.* 1957, Vol. 66, 4, pp. 177-182.
- 26 European Broadcasting Union. *Video player and recorder systems for home use* (2nd edition March 1972). Tech. 3093-E. Bruxelles, EBU, 1972 (Belgium)
- 27 Hadden, H.B. *High-quality sound production and reproduction. Chapter 11*. London, Iliffe, 1962.
- 28 Howard, J.A. and Ferguson, L.N. *Magnetic tape recording handbook*. Hewlett Packard application note No. 89, reprinted November 1970.
- 29 Ross, R.J. *Television film engineering*, Chapter 10, New York, Wiley, 1966.
- 30 Lunn, J.D. and Moffatt, M.E.B. Possible techniques for the recording of digital television signals. BBC Research Department Report No. 1969/42.
- 31 Jones, A.H. and Bellis, F.A. Digital stereo sound recorder. *Wireless World*, 1972, Vol. 78, 1443, pp. 432-435.
- 32 Bellis, F.A. An experimental digital television recorder. BBC Research Department Report No. 1976/7.
- 33 Jones, A.H. Digital television recording: a review of current developments. BBC Research Department Report No. 1973/29.
- 34 Hacking, K. Digital recording using hologram arrays: low-frequency micro-holograms on photographic film. See next article in this issue, pp. 21-39.
- 35 Bennett, W.R. and Davey, J.R. *Data transmission*, chapter 17, New York, McGraw-Hill, 1965.
- 36 Monteath, G.D. and Devereux, V.G. Digital methods in television. BBC Research Department Report No. EL 12, Serial No. 1967/55.
- 37 Croll, M.G. Pulse code modulation for high quality sound distribution: quantising distortion at very low signal levels. BBC Research Department Report No. 1970/18.
- 38 Devereux, V.G. and Meares, D.J. Pulse code modulation of video signals: subjective effect of random digit errors. BBC Research Department Report No. 1972/14.
- 39 Shorter, D.E.L., and Chew, J.R. Application of pulse code modulation to sound-signal distribution in a broadcasting network. *Proc. Instn. Elec. Engrs.* 1972, Vol. 119, 10, pp. 1442-1448.

- 40 Devereux, V.G. Pulse code modulation of video signals: subjective study of coding parameters. BBC Research Department Report No. 1972/40.
- 41 Bartolini, R., Hannan, W., Karlsons, D., and Lurie, M. Embossed hologram motion pictures for television playback. *Appl. Optics*, 1970, Vol. 9, 10, pp. 2283-2290.
- 42 Firester, A. H., Fox, E.C., Gayeski, T., Hannan, W.J., and Lurie, M. Redundant holograms. *RCA Rev.*, 1972, Vol. 33, 1, pp. 131-153.
- 43 Gerritsen, H.J., Hannan, W.J. and Ramberg, E.G. Elimination of speckle noise in holograms with redundancy. *Appl. Optics* 1968, Vol. 7, 11, pp. 2301-2311.
- 44 Vest, C.M., and Sweeney, D.W. Holographic interferometry of transparent objects with illumination derived from phase gratings. *Appl. Optics*, 1970, Vol. 9, 10, pp. 2321-2325.
- 45 Collier, R.J. and Pennington, K.S. Multicolour imaging from holograms formed on two-dimensional media. *Appl. Optics*, 1967, Vol. 6, 6, pp. 1091-1095.
- 46 Blackmer, L.L., van Kerkhove, A.P., Baldwin, R.E. and Stultz, K.F. Some aspects of multiple-beam interference techniques in digital data recording. *Appl. Optics* 1970, Vol. 9, 12, pp. 2753-2760.
- 47 Hitachi Ltd. Private Communication.
- 48 Gates, J.W.C. and Hall, R.G.N. Repetitive laser Q-switch. Proceedings of the Technical Program, Electro-optical systems design conference, Sept 22-24 1970, New York City, pp. 387-397. (Industrial-1 Scientific Conference Management, Inc. 1970.)
- 49 Macovski, A. Hologram information capacity. *J. Opt. Soc. Am.* 1970, Vol. 60, 1, pp. 21-29.
- 50 Langdon, R.M. Design of a large read-only holographic memory. *Radio and Electronic Engineer*, 1969, Vol. 38, 6, pp. 369-375.
- 51 Goodman, J.W. Introduction to Fourier Optics. pp. 62-63. New York, McGraw-Hill, 1968.
- 52 Leith, E.N., Upatnieks, J., Kozma, A. and Massey, N. Hologram Visual Displays. *J. Soc. Motion Pic. Tele. Engrs*, 1966, Vol. 75, 4, pp. 323-326.
- 53 van Heerden, P.J. Theory of optical information storage in solids. *Appl. Optics*, 1963, Vol. 2, 4, pp. 393-400.
- 54 Institution of Electrical Engineers. Permanent storage for holographic images. *Electronics and Power*. Vol. 17, October 1971, p. 380.
- 55 Childs, I. Laser beam deflection. BBC Research Department Report No. 1974/31.
- 56 Nelson, T.J. Digital Light Deflection. *Bell Syst. Tech. J.*, 1964, Vol. 43, 3, pp. 821-845.
- 57 Soref, R.A. and McMahon, D.H. Optical design of Wollaston-prism digital light deflectors. *Appl. Optics*, 1966, Vol. 5, 3, p. 425.
- 58 Steward, W.C., Mezrich, R.S., Cosentino, L.S., Nagle, E.M., Wendt, F.S. and Lohman, R.D. An experimental read-write holographic memory. *RCA Rev.*, 1973, Vol. 34, 1, pp. 3-44.
- 59 Cosentino, L.S. and Stewart, W.C. A membrane page composer. *RCA Rev.*, 1973, Vol. 34, 1, pp. 45-79.
- 60 Bobeck, A.H. The magnetic bubble. *Bell Labs. Rec.*, 1970, Vol. 48, 6, pp. 162-169.
- 61 Taylor, E.W. Improvements relating to information storage methods and systems. British patent No. 1,427,382.
- 62 Strull, G., List, W.F., Irwin, E.L. and Farnsworth, D.L. Solid-state array cameras. *Appl. Optics*, 1972, Vol. 11, 5, pp. 1032-1037.

Digital Recording using Hologram Arrays: Low-frequency Micro-holograms on Photographic Film

K. Hacking, B Sc.

Research Department

Summary: The article deals with the recording of micro-holograms, representing 'words' of binary-coded information, on conventional photographic film, and with the subsequent retrieval of this information. The work is the initial phase of a feasibility study of a possible system for recording digital television signals in real time using laser-beam scanning combined with holographic techniques.

A theoretical analysis of holographic image formation is followed by the results of basic experiments and measurements on film emulsions. The results show that binary storage densities of approximately 10^6 bits/cm² are feasible with average signal-to-interference ratios of 20dB or better after reconstruction, and with hologram spatial-frequency bandwidths of less than 200 cycles/mm. At this storage density or greater, recording and replay could be achieved using laser radiation powers of less than 10mW, assuming an optical efficiency greater than 10%.

- 1 Introduction
- 2 Theoretical considerations
 - 2.1 Fourier transform holography
 - 2.2 Analysis for a multi-aperture screen
 - 2.2.1 Imaging
 - 2.2.2 Recording
 - 2.2.3 Reconstruction
 - 2.3 Hologram efficiency
 - 2.4 Digital storage capacity
 - 2.5 Laser power requirements
 - 2.5.1 Writing
 - 2.5.2 Reading
- 3 Film sample measurements
 - 3.1 Transfer characteristic and sensitivity
 - 3.2 Modulation transfer function
- 4 Micro-hologram experiments
 - 4.1 5-bit holograms (40 μ m dia)
 - 4.2 9-bit holograms (30 μ m dia)
 - 4.3 Copying
- 5 Conclusions
- 6 References

- m = (s/f) = image magnification
 N = number of outer apertures
 P = storage density (bits/unit area)
 p = $2\pi/f\lambda$
 r = separation of outermost aperture from the reference aperture
 s = separation of reconstructing and hologram planes
 t_o = exposure time
 T_A = amplitude transmission coefficient
 W = laser radiation power
 U, V = rectangular co-ordinates referring to the reconstruction plane
 X, Y = rectangular co-ordinates referring to the screen plane
 x, y = rectangular co-ordinates referring to the recording plane
 β = operating (bias) point on film transfer characteristic
 λ = wavelength of radiation
 τ = hologram efficiency

List of Principal Symbols

- a = amplitude of plane coherent wave
 B = spatial-frequency bandwidth (cycles/mm)
 D = circular aperture diameter
 E_0 = average exposure required over hologram area
 ($1/E_0$ = film sensitivity)
 f = focal length of lens
 g = $2\pi/\delta\lambda$
 h = separation of adjacent apertures in screen
 j = $\sqrt{-1}$

1 Introduction

Over the past few years there has been an intensive search in various laboratories¹⁻³ for new recording media and techniques capable of storing vast quantities of binary data with rapid access, fast transfer rates and sophisticated updating facilities, all (needless to say) at low cost. Many of these requirements and ideals are common to the problem of recording digital television signals in real time, and a review of the possibilities (prior to 1969) was given in a Research Department Report.⁴ Subsequently, a growing interest has been shown in optical storage for binary data using

holographic techniques, particularly in computer applications.⁵ The concept which has emerged involves arrays of small holograms addressed by a laser beam, the reconstructed image so formed being read by a matrix of photo-sensors. Typically, for a read-only optical store, each hologram would be about 1mm in diameter and contain 10^6 bits.

Redundancy is an intrinsic feature of holographic records which can be exploited, although not always easily. The general principle is that the information relating to each point of the original scene is spread⁶ over the whole area of the hologram. Thus a small localised defect, e.g. a dust particle or scratch, on the record would have only a small effect on the reconstructed image. This advantage of defect immunity, combined with the ruggedness achieved by binary coding, potentially offers a reliable and attractive system of digital data storage with high packing density.

During 1970 various methods were suggested for applying these basic ideas of holographic storage to the real-time recording of digital television signals.⁷ Although the systems proposed were elegant in concept, their practical implementation presented formidable problems in terms of existing technology; large arrays of fast-acting light valves and photo-sensors were required and such devices have yet to be developed to the stage of practicability. As an introductory step, an alternative system⁸ was devised using micro-hologram arrays formed on conventional photographic film by a laser scanning arrangement. It was envisaged initially that each hologram would contain only a single word of eight or nine bits and have a diameter of approximately $30\mu\text{m}$ (roughly equivalent to the diameter of a television picture element in conventional analogue recording on 35mm cine film). Although any hologram array format may be used in principle, there would be a possible advantage in a line-by-line array, with a conventional aspect

ratio, because a slight modification of the system might then permit a faint analogue version of the television picture to be recorded simultaneously. Such a picture, overlaid on the digital recording, could be useful for visual identification or editing.

A disadvantage of using micro-holograms is that, for a given packing density, a great deal, but not all, of the intrinsic holographic advantage of blemish immunity is sacrificed: this loss can be mitigated, however, by introducing redundancy in the binary coding. There are, however, two important practical advantages:—

- 1 The number of fast-acting photo-sensors and light valves required is dramatically reduced, so that discrete transducers become practicable (and are, in fact, available).
- 2 Because each micro-hologram contains only a few bits of information, an axial reference beam⁸ can be used (see section 2.1) with the result that the hologram spatial-frequency bandwidth can be greatly reduced, in fact to less than 200 cycles/mm.

The reduced bandwidth not only permits less critical physical tolerances during recording, and lower numerical lens apertures, but also reduces trouble from component surface dust and speckle effects. Another practical advantage is that higher-sensitivity photographic emulsions can be used, with a consequent reduction in the power of the laser source.

The work reported here is the initial phase of a feasibility study aimed at establishing a working system for recording digital television signals by means of hologram arrays. It is principally concerned with an investigation of low-frequency micro-holograms formed on several commercially-available photographic emulsions. The work has involved the design and construction of a laser-source test bench and a camera head, as well as the development of various measurement techniques. Measurements of

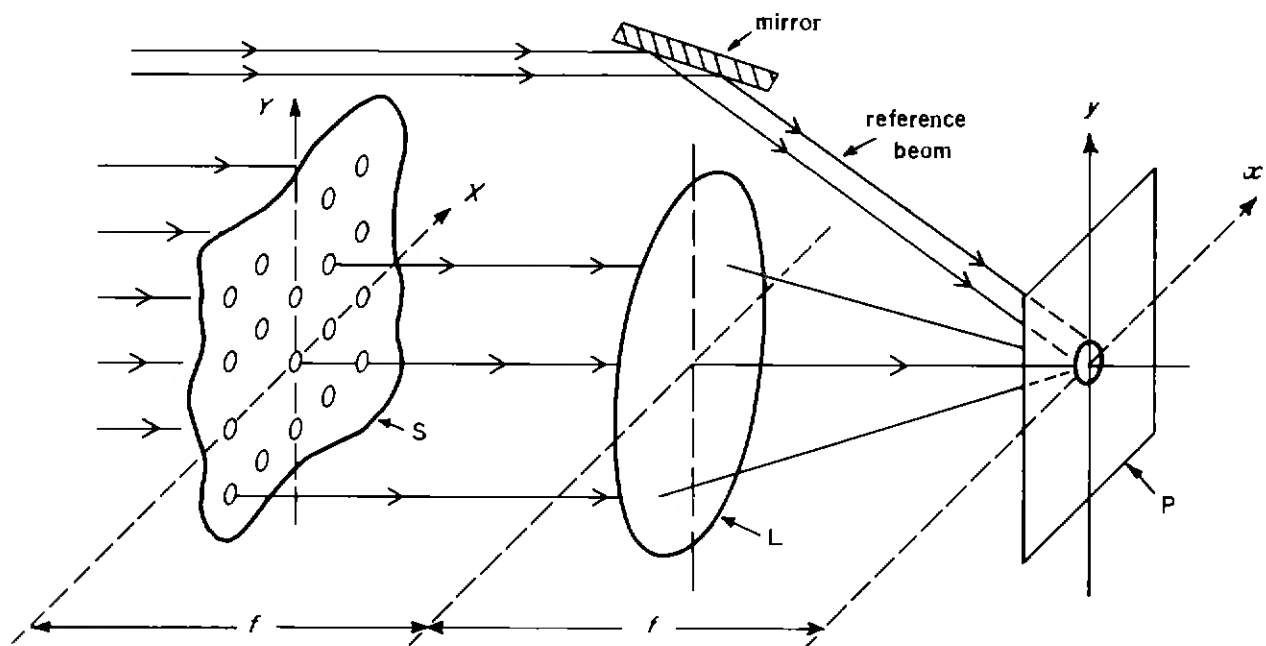


Fig. 1 Typical arrangement for FT hologram recording

resolution and transfer characteristics have been carried out on three different emulsions. These results, supported by data from other laboratories, are used to assess the maximum efficiency which can be expected and the relative importance of several fundamental parameters.

2 Theoretical considerations

2.1 Fourier transform holography

A typical arrangement for producing Fourier transform (FT) holograms is shown in figure 1. The transparency whose hologram is to be recorded is placed in the front focal plane (X, Y) of a lens, L . As indicated in figure 1, the transparency could be a thin screen, S , with a matrix of small holes in it representing 'bits' of information (the presence or absence of a hole at a given point defining the binary state). A photographic plate or film, P , is placed in the rear focal plane (x, y) of the lens. During an exposure, the aperture screen is illuminated, as shown, by a collimated beam of coherent radiation.* For those readers not familiar with lens systems, figure 1 might give the erroneous impression that P receives an actual image of the screen S . In fact, the light emerging from the lens is concentrated in a single, small spot at P .

* The purpose of the aperture screen, when rear-illuminated, is to generate a set of coherent beams, sometimes called 'digit' beams. The insertion of an actual screen with apertures is not essential and other⁹, perhaps more suitable, methods of generating the coherent beams can be used. The treatment here only requires that the complex amplitude distribution of the radiation in the screen plane be specified.

The situation can be described more concisely in mathematical terms. If $F(X, Y)$ is the complex amplitude distribution of the radiation in the screen plane, and $G(x, y)$ is the corresponding distribution in the recording plane, then $G(x, y)$ is the Fourier transform of $F(X, Y)$, assuming the lens is perfectly corrected for geometrical aberrations. Hence, by adding simultaneously an inclined reference beam and then exposing the plate, a carrier-type hologram of the Fourier transform of the hole pattern in the screen can be formed.

The proposals and experiments described later in this article are based on an alternative arrangement in which the reference beam is axial rather than inclined. An axial reference beam can be derived from a central aperture in the binary digit screen. The new arrangement is shown in figure 2, where the central (reference) aperture of the screen is surrounded by only a small number of outer apertures which could, for example, specify a single binary word. As will be shown later, the relative positions of the outer holes are carefully selected to avoid certain cross-talk effects appearing in the reconstruction. This method of holography results in a baseband-type hologram having a relatively small spatial-frequency bandwidth.

2.2 Analysis for a multi-aperture screen

The principal features of the imaging, recording, and reconstruction processes in FT holography can be gleaned from an analysis assuming an elementary, multi-aperture object screen.

2.2.1 Imaging

Let the object, or digit screen, consist of an arbitrary array of N outer apertures located in the X, Y plane at the co-ordinates $X_1 Y_1, X_2 Y_2 \dots X_N Y_N$, as indicated in figure 3a,

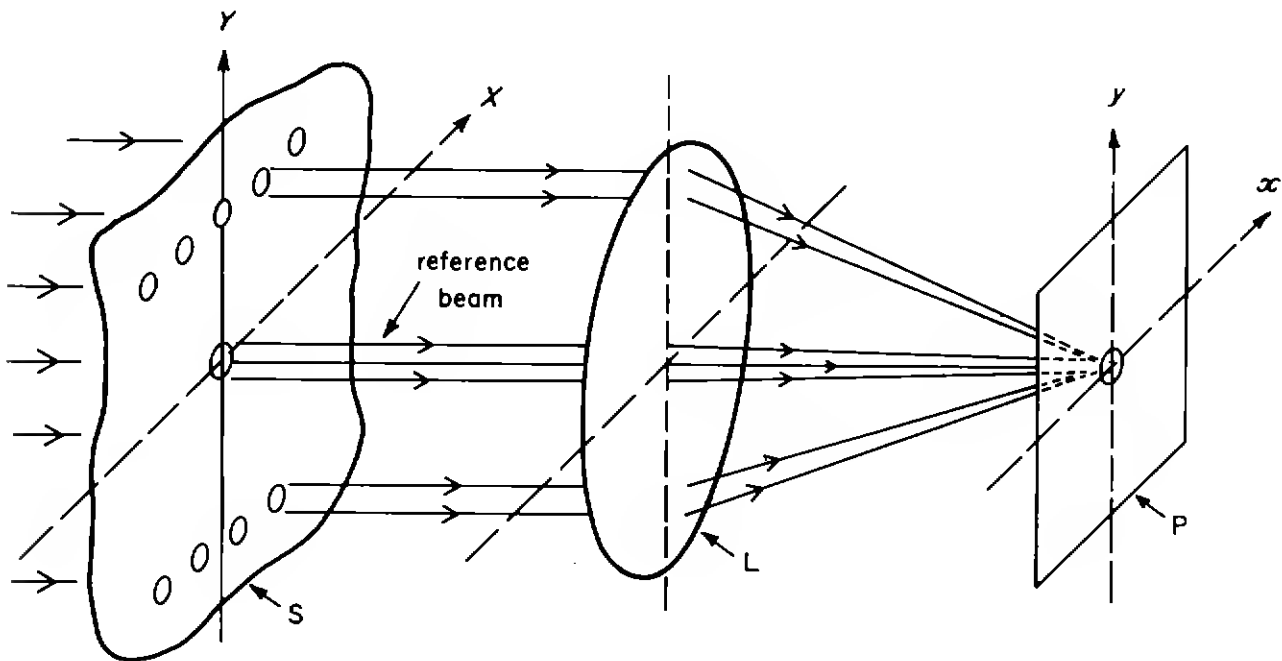


Fig. 2 FT hologram recording using an axial reference beam

and a central reference aperture at $X_0 Y_0$. All the apertures, including the reference, are assumed to be identical in size and form. During exposure, the apertures are illuminated by plane wave-fronts derived from a common coherent source, but with field amplitudes a_0, a_1, \dots, a_N at the respective apertures. Further let the function $F(X, Y)$ describe the spatial variation of the amplitude of the wave transmitted by each aperture.

It can be shown (using the displacement theorem of Fourier transforms) that the total amplitude distribution $A(x, y)$ in the recording plane (figure 3b) is given by

$$A(x, y) = G(x, y) \sum_{n=0}^{n=N} a_n \exp [-jp(X_n x - Y_n y)] \quad (1)$$

$$\text{where } G(x, y) = \iint_Z F(X, Y) \exp [-jp(Xx - Yy)] dX dY$$

$p = 2\pi/\lambda$, f = focal length of lens, λ = wavelength of radiation, Z = aperture zone, $F(X, Y) = 0$ outside Z , n = integer ($n = 0$ refers to the central reference aperture).

The spatial variation in the intensity of the radiation falling on a photographic plate, placed in the recording plane x, y , is equal to the square of the modulus of $A(x, y)$. Hence the exposure $E(x, y)$ is given by

$$E(x, y) = t_0 A(x, y) A^*(x, y)$$

where t_0 = exposure time and $*$ denotes the complex conjugate. Thus, substituting from equation 1,

$$E(x, y) = t_0 |G(x, y)|^2 \left\{ \sum_{n=0}^{n=N} a_n \exp [-jp(X_n x - Y_n y)] \right\} \left\{ \sum_{n=0}^{n=N} a_n^* \exp [jp(X_n x - Y_n y)] \right\} \quad (2)$$

Equation 2 represents a complicated multi-beam interference pattern confined within the Fraunhofer diffraction pattern of a single aperture. The maximum spatial frequency within the total pattern is determined by the greatest spacing between any two apertures in the screen, the focal length of the lens, and the wavelength of the radiation.

2.2.2 Recording

After exposure, the photographic plate or film is processed in the conventional manner to form a permanent hologram record. Alternatively, the silver image can subsequently be bleached out to form a non-absorbing phase hologram. Using non-bleached holograms, the processed record should ideally have an amplitude transmission factor, T_A , that is

linearly related to the spatial variation of intensity during exposure.¹⁰ That is

$$T_A(x, y) \propto (1/t_0) E(x, y) \quad (3)$$

Suppose we consider, however, using a negative hologram for the reconstruction, the negative having the transfer characteristic shown in figure 4a. Here the development process is specified by the relation between the amplitude transmission factor of the negative record and the normalised exposure. The dashed line represents the ideal linear case, while the full line is typical of the relation achieved with high-resolution film emulsions. For comparison, figure 4b shows the same ideal and practical data replotted in the more conventional (H and D) form of optical density versus the logarithm of the exposure. Continuing the analysis for the linear but negative case, we obtain instead of the proportionality 3

$$T_A(x, y) = 1 - \beta [E(x, y)/E_0] \quad (4)$$

Where E_0 is the average exposure of the hologram. The corresponding average value of T_A is $1 - \beta$, where β is an arbitrary operating (bias) point on the transfer characteristic of the recording emulsion selected by adjusting the average exposure (see figure 4a). Two important general points concerning the hologram formation and the film exposure become apparent from figure 4 and equation 4. These are

- 1 the point-to-point variations of intensity in the light forming the hologram should be confined to as small a range of intensity as possible in order to mitigate the effects of non-linearity. (Even in the ideal linear case we require that $E(x, y)$ be always less than $1/\beta$ times its average value.)

- 2 there is not a great latitude in the bias point, β , for a practical system, without significantly changing the degree of transfer non-linearity. The actual tolerances are best determined by experiment.

2.2.3 Reconstruction

After processing, the negative hologram is replaced in its original position in the x, y plane and illuminated only by the reference beam (i.e. with only the central aperture of the object screen open) as shown in figure 5. The reconstructed image of the original outer apertures may be examined by placing a viewing screen (of ground glass, say) in a plane (U, V) at a distance s from the hologram plane. It is assumed in the following that the distance s is very large compared with the hologram dimensions.

The spatial variation of amplitude of the light emerging

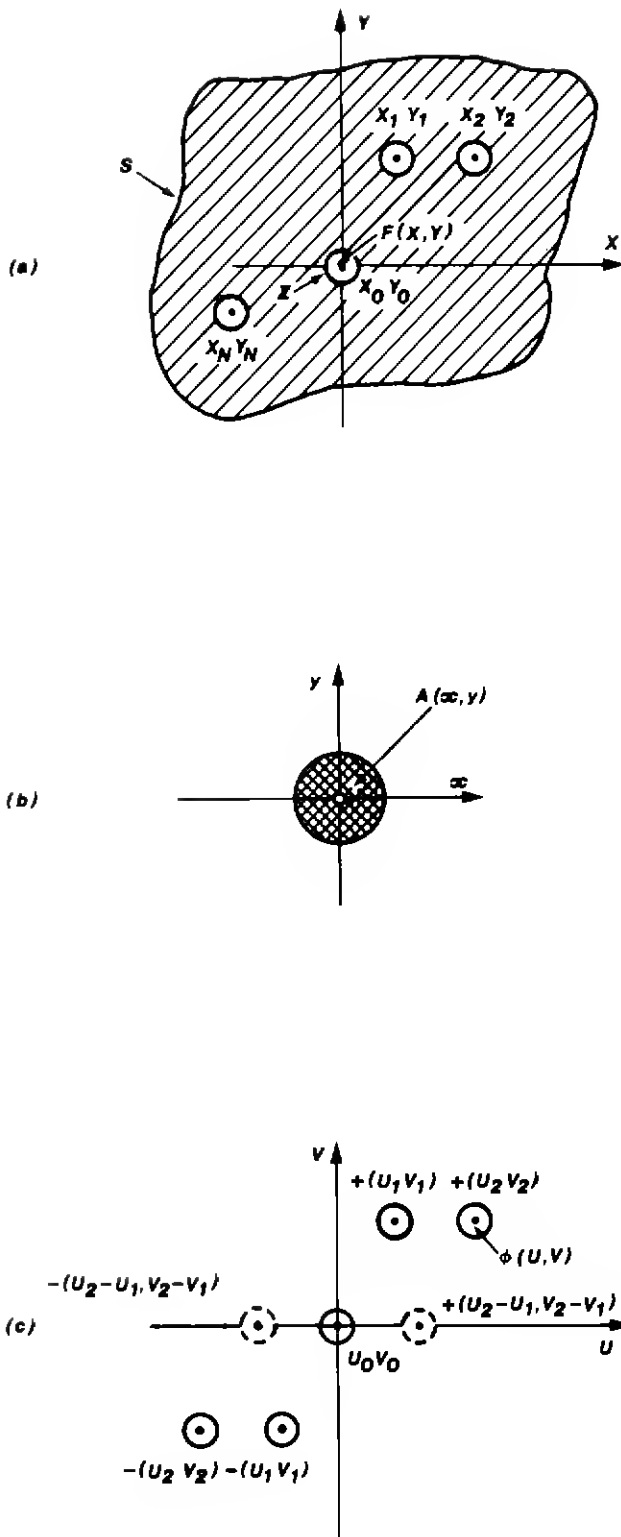


Fig. 3 Analysis notation for a multi-aperture object screen
 a. Object screen with N outer apertures
 b. Amplitude distribution in Fourier Transform (hologram recording) plane
 c. Reconstruction plane and image set for $N = 2$.

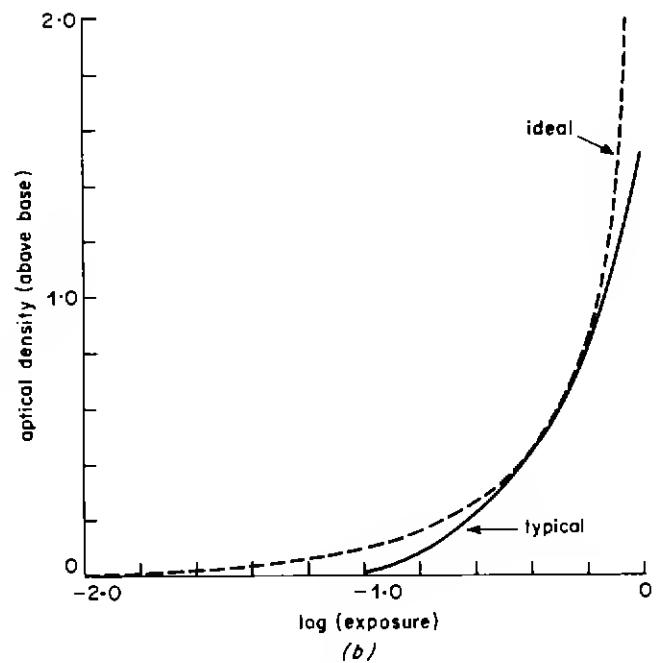
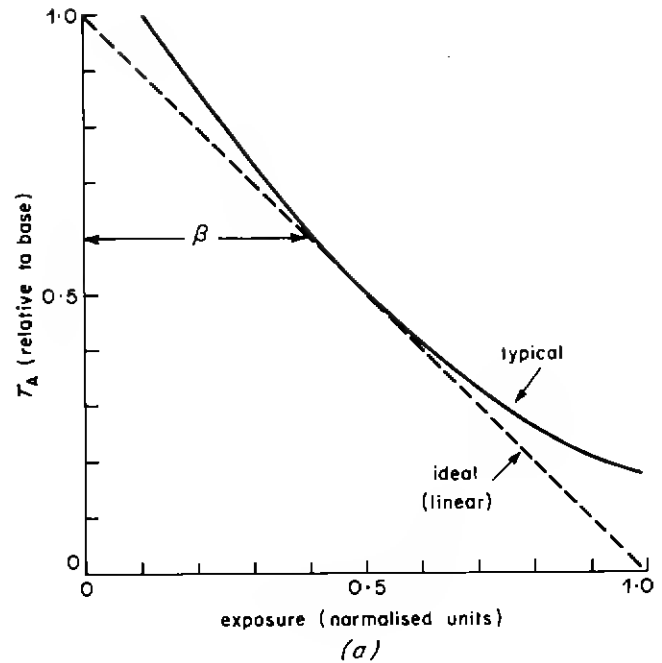


Fig. 4 Negative film transfer characteristics for hologram recording
 a. Amplitude transmission factor T_A , as a function of relative exposure
 b. Conventional H and D plot of same data

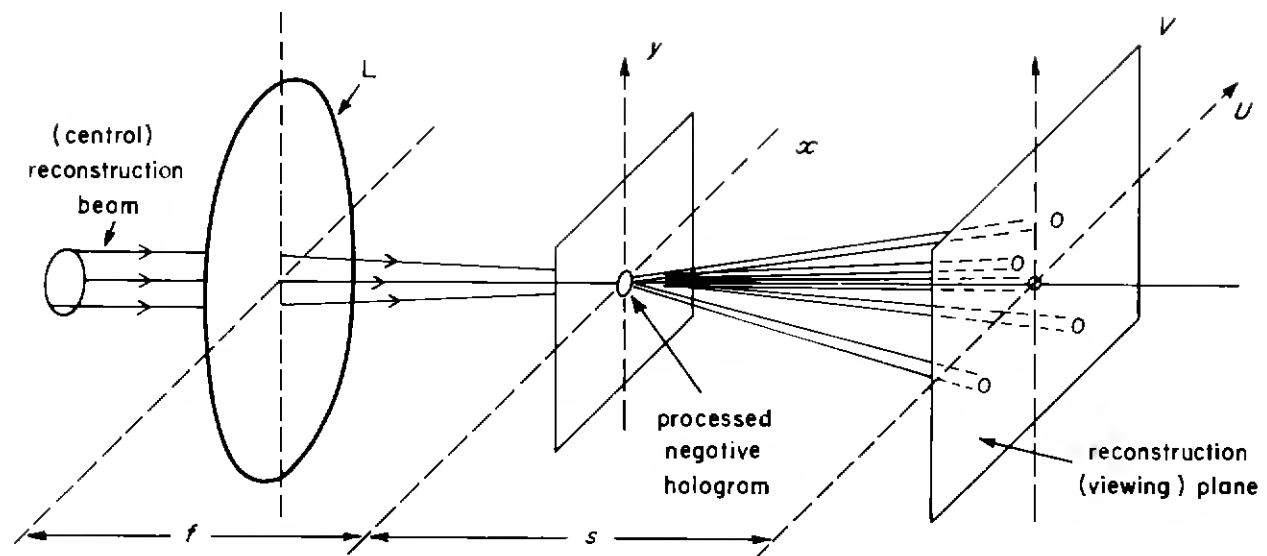


Fig. 5 Reconstruction notation for FT hologram

from the negative hologram, $H(x,y)$ say, is

$$H(x,y) = T_A(x,y)G(x,y) \quad (5)$$

which, using equation 4, becomes

$$H(x,y) = G(x,y) - (\beta/E_0)G(x,y)E(x,y) \quad (6)$$

Finally, the (far-field) diffraction pattern, $I(U,V)$ falling on the viewing screen S to form the reconstructed image in the U,V plane (see figure 5) is derived from the two-dimensional Fourier transform of $H(x,y)$. Thus

$$I(U,V) = \left| \iint_{-\infty}^{\infty} H(x,y) \exp[jg(xU + yV)] dx dy \right|^2 \quad (7)$$

where $g = 2\pi/s\lambda$.

Although equations 5 and 7 formally describe the process for a multi-aperture object screen, the characteristics of the reconstructed image are more easily seen by looking at the individual terms obtained for a simple object screen with two outer apertures, say. Thus putting $N = 2$ in equation 2 and using equations 6 and 7, one obtains:

$$I(U,V) = \left| F(U,V) - \beta\phi(U,V) - \frac{\beta a_0 a_1}{(a_0^2 + a_1^2 + a_2^2)} \phi(U \pm mX_1, V \pm mY_1) - \frac{\beta a_0 a_2}{(a_0^2 + a_1^2 + a_2^2)} \phi(U \pm mX_2, V \pm mY_2) - \frac{\beta a_1 a_2}{(a_0^2 + a_1^2 + a_2^2)} \phi[U \pm m(X_2 - X_1), V \pm m(Y_2 - Y_1)] \right|^2 \quad (8)$$

where $m = s/f$ (image magnification) and $\phi(U,V)$ is a normalised and scaled Fourier transform of the function $G(x,y) = |G(x,y)|^2$. In fact, $\phi(U,V)$ can be regarded as a blurred version of the aperture function $F(U,V)$ ($\equiv F(X,Y)$ except for change of variable).

It may now be seen, from equation 8, that the reconstruction of an object screen with $N = 2$ consists of six separate light distributions having the form $\phi(U,V)$ and

grouped into three sets of identical twin images. Each twin set has a different peak intensity. The images are located at the points $\pm(mX_1, mY_1)$, $\pm(mX_2, mY_2)$ and $\pm[m(X_2 - X_1), m(Y_2 - Y_1)]$ respectively. In addition a different distribution, $F(U,V) - \beta\phi(U,V)$, having the largest peak intensity, occurs on the optical axis, i.e. at $U = 0, V = 0$.

The reconstructed version of a two-aperture screen (plus the central reference aperture) is shown in figure 3c. In addition to those images having the same geometrical disposition as the original apertures, there is a duplicate set mirrored in the opposite quadrant, plus a spurious pair of images; these latter images are symmetrically disposed about the origin and are separated from the origin by an amount equal (or proportional) to the original separation of the screen apertures.

2.3 Hologram efficiency

Small photo-detectors placed in the reconstruction plane at the corresponding geometrical positions of the N original screen apertures can be used to indicate whether or not a

particular aperture was 'open' during the exposure of the spot hologram. This provides a means of recovering the original binary word in electrical form.

Now, if the (blurred) images in the reconstruction are sufficiently well separated, the analysis in section 2.2 above suggests that the intensity distribution of the n th wanted image in the reconstruction is given by an expression of the form

$$\left[\beta a_n a_o \left(\sum_{n=0}^N a_n^2 \right)^{-1} \right]^2 \left| \phi(U, V) \right|^2$$

The factor in square brackets indicates the theoretical efficiency of the process while the other factor describes the spatial variation of image intensity. Both are important in relation to the maximum storage capacity of the individual holograms and the signal-to-interference ratio of the system.

The hologram efficiency, τ , can be defined as the ratio of the average power of the radiation forming a wanted image to that of the reconstructing beam. For a non-bleached hologram, having a transfer characteristic specified by equation 4 and with perfect resolution characteristics, the hologram efficiency for the n th aperture of a multi-aperture screen is

$$\tau = \left[\beta a_n a_o \left(\sum_{n=0}^N a_n^2 \right)^{-1} \right]^2 \quad (9)$$

where it is assumed that the N outer apertures and the reference aperture have identical size and form. If it is assumed that all the outer apertures were illuminated with equal intensity during the exposure, equation 9 reduces to

$$\tau = \frac{\beta^2 (a/a_o)^2}{[1 + N(a/a_o)^2]^2} \quad (10)$$

where a^2 = illumination intensity of the N outer apertures.

Thus, in this case, the hologram efficiency depends on

- (i) the bias-point, β , on the transfer characteristic,
- (ii) the number N of outer (signal) apertures,
- (iii) the ratio $(a/a_o)^2$, which is the ratio of the illumination intensities of the outer apertures to that of the reference aperture.

It is easily shown that the maximum value of τ in equation 10 is obtained when $N(a/a_o)^2 = 1$, i.e. when the sum of the illumination intensities of the N outer apertures equals that of the reference aperture. Thus, from equation 10,

$$\tau_{\max} = \frac{\beta^2}{4N} \quad (11)$$

Figure 6a shows the variation of the hologram efficiency, τ , with $(a/a_o)^2$ for $N=5$ and for $N=9$, as given by equation 10 with $\beta=\frac{1}{2}$. Figure 6b shows the inverse proportionality of τ_{\max} with N , as given by equation 11, again with $\beta=\frac{1}{2}$.

The modulation transfer functions or resolution characteristics of actual photographic emulsions tend to droop with increasing spatial frequency, and this reduces the efficiency of the holographic process. Equalisation of the efficiencies of the individual apertures can be achieved, however, by a compensatory adjustment of their respective illumination intensities at the time of exposure.

Variations in the hologram efficiency can also be expected with deviations from linearity in the negative transfer

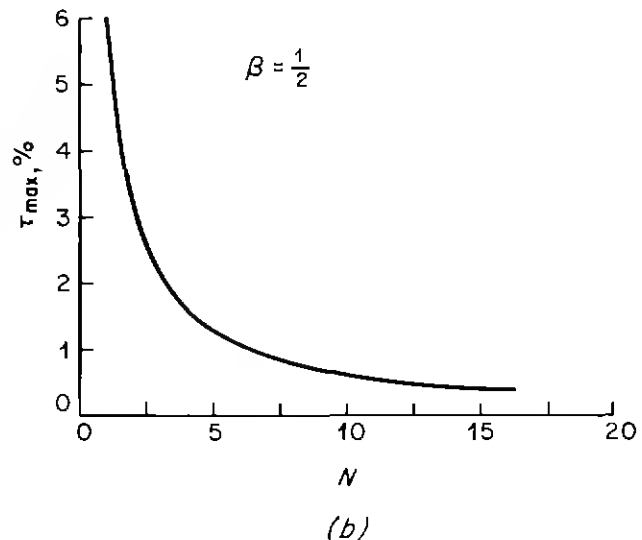
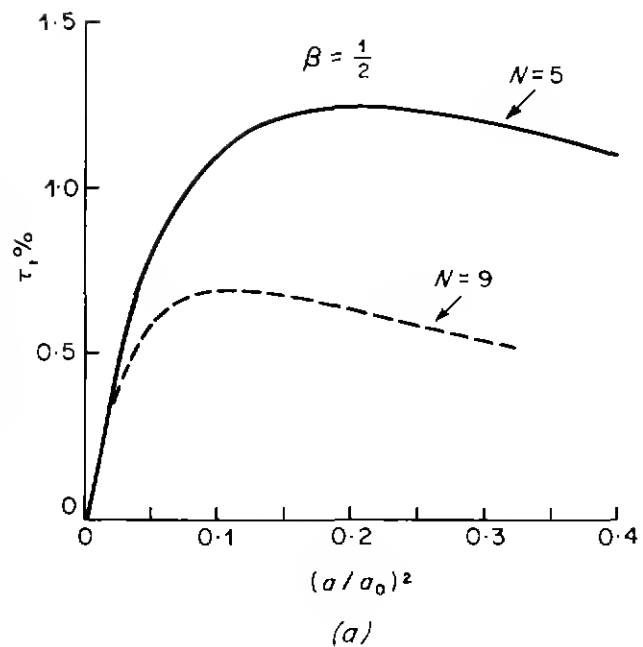


Fig. 6 Theoretical efficiency for a non-bleached hologram of a binary screen with N identical outer apertures

- a. Variation of efficiency with ratio of outer to reference aperture illumination for $N=5$ and $N=9$
- b. Maximum efficiency of process as a function of number of outer apertures

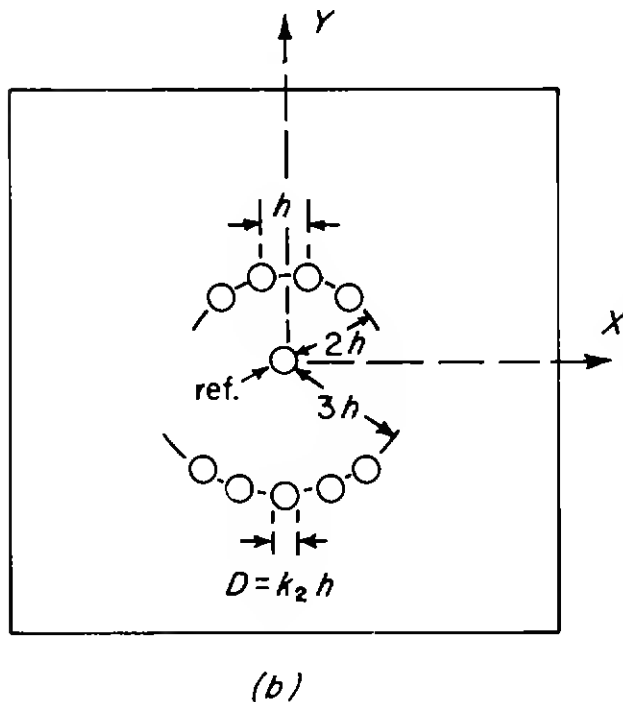
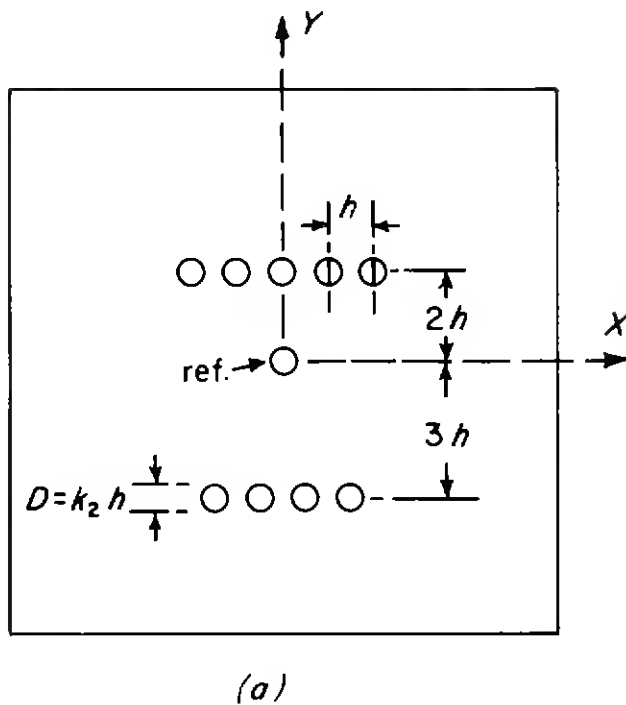


Fig. 7 Two types of binary aperture screen for low-bandwidth FT holograms

characteristic and from changes in the bias point.

Greater efficiencies are possible by bleaching the holograms after development to produce non-absorbing phase holograms, although some adverse features have been reported.¹¹

2.4 Digital storage capacity

Using the concept that each separate aperture position (except the reference aperture) in the object screen corresponds to one 'bit' of information, the storage per hologram spot is simply equal to the number N of such positions. If Z is the effective area of the reference aperture, the corresponding area of the hologram spot is proportional to $f^2 \lambda^2 Z^{-1}$ where, as previously, f is the focal length of the lens and λ is the wavelength of the light. (This relationship arises due to diffraction at the aperture.) Hence, thinking in terms of a two-dimensional array of small holograms, the storage density P (bits/unit area) is

$$P = \frac{k_1^2 N Z}{f^2 \lambda^2} \quad (12)$$

where k_1 is a coefficient determined by the closeness with which the holograms can be packed in the array without producing significant cross-talk between neighbouring holograms. The value of k_1 is best determined by experiment, and is expected to be less than unity.

The achievable storage density must, however, be weighed against other basic criteria such as the overall error rate in the reconstituted binary signal.

The concept of micro-hologram arrays of single binary words leads to a holographic system of low spatial-frequency bandwidth because the low number of bits per hologram allows an axial reference beam to be used. In this case, the spatial-frequency bandwidth of the system is given by $r/f\lambda$, where r is the separation in the plane of the aperture screen of the outermost aperture from the reference aperture.* Suppose we have an aperture screen representing a 9-bit word arranged as shown in figure 7a. The circular apertures lie along two rows, $Y = 2h$ and $Y = -3h$ respectively. The apertures have a diameter D and are spaced h apart ($h > D$). It can easily be seen that, with this design, the spurious images arising in the reconstruction, due to a simultaneous exposure of all the apertures, lie clear of the wanted set of images. Moreover, any second order diffraction images arising from transfer non-linearity will also lie clear of the wanted set of images. In an alternative design, the apertures lie on circular arcs of radii $2h$ and $3h$ respectively as shown in figure 7b. For this particular aperture screen arrangement, r is equal to $3h$ and therefore the system spatial-frequency bandwidth is $3h/f\lambda$. In order to avoid significant crosstalk between neighbouring images in

*Two parallel coherent beams separated by a distance r and brought to a focus by a lens of focal length, f , produce a sinusoidal variation of intensity in the focal plane which has a (maximum) spatial frequency equal to $r/f\lambda$, where λ = wavelength of the radiation.

the reconstruction the screen aperture diameters must be somewhat less than their spacing, i.e. $D = k_2 h$ where k_2 is a coefficient less than unity. Thus, in terms of the spatial-frequency bandwidth, B , the storage density using the nine-aperture screen shown in figure 7b becomes (substituting for $f\lambda$ and NZ in equation 12)

$$P = (k_1 k_2 B)^2. \quad (13)$$

As a numerical example, setting $k_1 = k_2 = 1/\sqrt{2}$ (which is about the practical limit) and $B = 200$ cycles/mm, the storage density is 10^4 bits/mm² (10^6 bits/cm²). At this density, a high-quality digital television recording of a scene would occupy an area of film equal to that of a conventional 35mm cine-film frame.

An estimate of the minimum spacing of the apertures in the object screen, i.e. of the value of k_2 in equation 13, can be obtained by calculating the reconstructed image profile for a single aperture. It has already been mentioned (section 2.2.3) that the reconstructed image will be a blurred version of the original aperture. The image profile ($\phi(U, V)$ in equation 8) calculated for the case of a circular aperture is shown in figure 8. It will be seen that the profile spreads to approximately twice the original aperture diameter, so that for no interference between adjacent images the minimum aperture separation is also this value, i.e. $k_2 = \frac{1}{2}$. However, some overlap of the images is permissible in practice, especially if only the core of the image is used for binary on/off sensing.

Although the salient features of the system can be deduced by considering circular apertures uniformly illuminated, there are alternative variations which could lead to less crosstalk, not only between adjacent reconstructed images, but also between adjacent holograms in the recorded array. For example, by tailoring the radial variation of amplitude and phase of the radiation leaving the primary screen apertures (i.e. by 'apodisation') the hologram spot could be more uniformly exposed. This would be especially beneficial because of the small latitude for correct exposure. However, a full investigation of these possible refinements is beyond the scope of this article.

2.5 Laser power requirements

In this section, the factors which control the input and output data transfer rates are considered in general terms. The transfer rate required for real-time television recording and replay in digital form is of the order 10^8 bits per second, and it is this very high rate which really stretches the existing technology with regard to suitable transducers.

2.5.1 Writing

The basic factors which determine the potential transfer rate in recording hologram arrays on photographic film are

- (i) sensitivity of the film
- (ii) modulation transfer function (MTF) of the film
- (iii) optical efficiency of the system

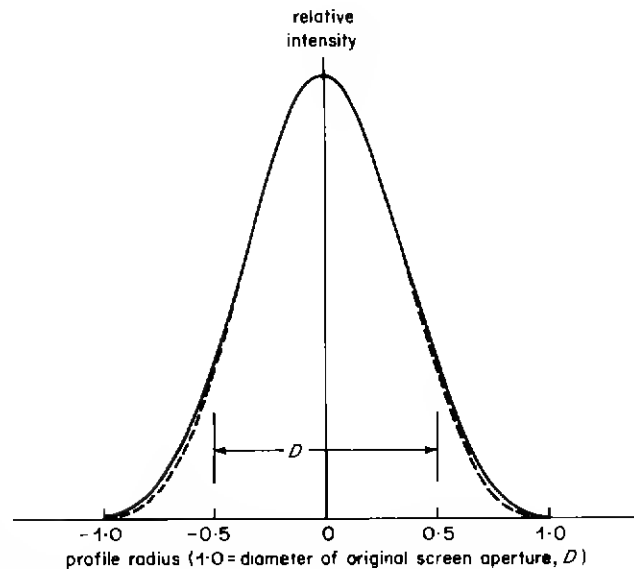


Fig. 8 Calculated intensity profile (principal section) of reconstructed image

— linear transfer characteristic
 --- typical transfer characteristic

(iv) laser radiation power.

It is well known that the first two factors are somewhat incompatible in the sense that high sensitivity and high resolution characteristics do not go together. Many of the emulsions used for conventional holography have very high resolution capabilities (>2000 cycles/mm) but are also exceedingly slow.

The third factor, optical efficiency, depends on the particular system of recording and scanning devised. It might be thought that high efficiencies could be obtained easily with coherent laser beams, but this is not so, especially if sophisticated beam-deflection devices and fast modulators are required. Overall optical efficiencies greater than about 20% could be difficult to achieve, at least in the early stages of development.

The laser radiation power, W , required for real-time recording at an input transfer rate of I bits/sec is given by

$$W = \frac{IE_0}{\mu P} \times 10^{-4} \text{ (mW)} \quad (14)$$

where μ is the overall optical efficiency of the system, P is the information packing density (bits/cm²) and $1/E_0$ is the reciprocal exposure or sensitivity of the recording film in ($\mu\text{J}/\text{cm}^2$)⁻¹.

Figure 9 shows a plot of laser power as a function of packing density with film sensitivity as parameter; a transfer rate of 100 Megabits/sec and an optical efficiency of 10% have been assumed. The film sensitivity values cover the range of commercially-available holographic emulsions. For micro-hologram arrays, the packing density is proportional to the square of the system bandwidth (see equation 13), and the modulation transfer factor of the recording film must be adequate at the maximum spatial frequency required. Furthermore, on substituting equation 13 in equation 14 and then re-arranging, the product $(1/E_0)B^2$ occurs in the denominator of the expression for the laser power. Now, for a

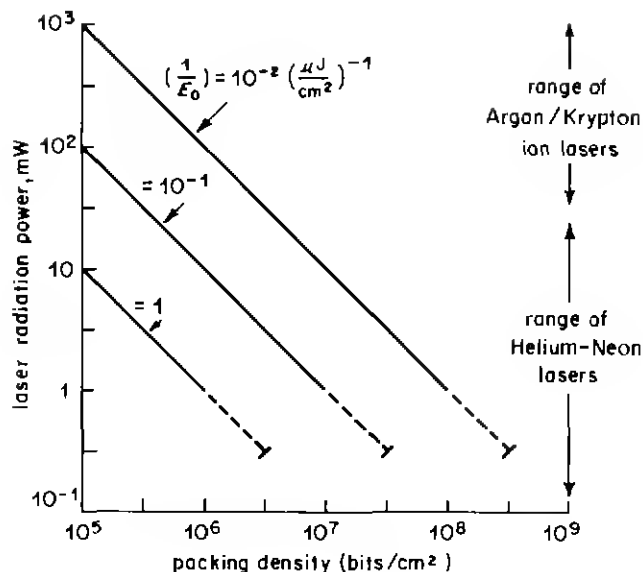


Fig. 9 Laser power required for holographic recording as a function of packing density with film sensitivity ($1/E_0$) as parameter
Optical efficiency = 10%; recording rate = 10^8 bits/sec

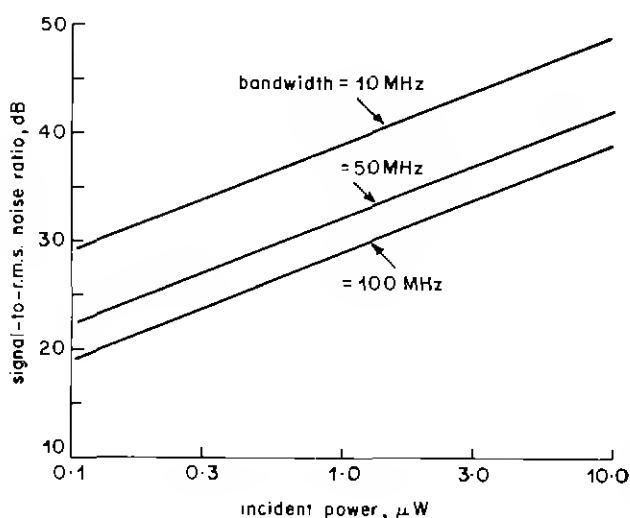


Fig. 10 Limiting signal-to-noise ratios for a photo-detector (quantum efficiency = 5%) as a function of incident power, with detector bandwidth as parameter

range of recording films the product of the sensitivity and the square of the spatial bandwidth tends to be roughly constant, and this is the reason that the limit marks in figure 9 correspond to a constant power. The diagram illustrates that the slower but higher-resolution films are necessary for television recording at very high packing densities.

Providing that an optical efficiency in the recording arrangement greater than 10% can be achieved, it would appear that low-power gas lasers (5 to 10 milliwatts) would be suitable for a television recording system using micro-hologram arrays.

2.5.2 Reading

The principal criterion in relation to the laser power required for the reconstruction or read-out system is that the overall error rate is not significantly increased by photon noise in the reading beam. The limiting values of signal-to-noise ratio as a function of radiation power at $\lambda = 632.8\text{nm}$ (He-Ne) incident on a photodetector are shown in figure 10, for various values of detector bandwidth. A photodetector quantum efficiency of 5% has been assumed (typical value for an S20 photo-cathode at this wavelength).

The reconstruction efficiency to be expected with non-bleached holograms was discussed earlier (section 2.3), and values between 0.1% and 0.5% are typical. If we regard as adequate a limiting signal-to-noise ratio (due to photon noise) of 40dB with a detector bandwidth of 50MHz, then we require an incident radiant power of at least $6\mu\text{W}$ on each detector (5% quantum efficiency). The required radiation power incident on each hologram, therefore, falls in the range 1.2mW to 5mW in this example. Assuming an optical efficiency of 20% in the read-out apparatus, the laser power required is at least 6mW, which is commensurate with that required for writing.

The laser power required for reading could be significantly reduced by using a photodetector with a high quantum efficiency at $\lambda = 632.8\text{nm}$, e.g. a fast silicon photodiode having a quantum efficiency of 60% or greater.

3 Film sample measurements

Some basic measurements were made on several commercially-available film emulsions which were thought to be potentially useful for low-frequency micro-hologram recording. It appeared originally, from manufacturers' data, that there was a considerable range of film emulsions available. However, the actual selection available off-the-shelf was limited, largely due to recent 'rationalisation' of these materials by the principal manufacturers. In the event, four film stocks (Ilford Micro-Neg Pan; Agfa Copex Pan Rapid; Kodak Microfile 5669; Agfa Scientia 10E 75) were investigated experimentally. After the initial measurements it was found that Agfa Copex Pan Rapid was similar to Ilford Micro-Neg Pan and, therefore, was not included in the subsequent experiments.

3.1 Transfer characteristic and sensitivity

The transfer characteristics were measured by exposing the emulsions through a calibrated neutral-density step wedge to a uniform beam of laser radiation (He-Ne laser at $\lambda = 632.8\text{nm}$). The results are shown in figure 11, where the amplitude transmission coefficient is plotted as a function of exposure (log scale). Throughout the work the development conditions were held constant; 4 minutes in Kodak D.19 (or Ilford Phentrace) at 20°C with nitrogen bubble-burst agitation every 30 seconds. The exposure times were in the region of 40 ms, i.e. near the maximum photometric efficiency.

Sensitivity, for this application, may be defined as the reciprocal of the exposure required to reach a given point of the transfer characteristic. A convenient point is where the amplitude transmission coefficient equals $\frac{1}{2}$, for which an exposure E_0 is required; these bias points are shown on the curves in figure 11.

The results shown in figure 11 refer to relatively long exposures, and it is well known that for much longer or much shorter exposure times reciprocity fails; the effect is to diminish the photometric efficiency of the emulsion. Measurements made on the film samples indicated a decrease in effective sensitivity by a factor of 3 for an exposure time of about 100 microseconds ($\times 3$ reciprocity factor). The experience of other laboratories working on pulsed holography is that for exposure times below about 10 micro-

seconds there is no further loss of sensitivity. For exposures in the nanosecond range, the photometric efficiency of a holographic-type emulsion is expected to be not less than one quarter of its maximum value.

3.2 Modulation transfer function

The MTF of film emulsions at spatial frequencies greater than 100 cycles/mm is not easily measured by conventional methods, e.g. by scanning line-spread functions or other recorded image patterns. A more satisfactory approach is to record a sine-wave intensity pattern, directly generated by two interfering beams of coherent light, and then to measure the diffraction efficiency of the processed record, again using a coherent light beam.

Figure 12a shows the arrangement used to generate sets of interference fringes of various fixed spatial frequencies, with control of modulation depth. Figure 12b indicates the method used for measuring the diffraction efficiency of the test recordings.

The principal results obtained are shown in figure 13, where each curve plotted is the variation of the intensity of the first-order diffraction component (n_1) as a function of the zero-order intensity (n_0), for the stated spatial frequency and for 100% modulation depth of the interference fringes during exposure. The intensities are expressed as a per-

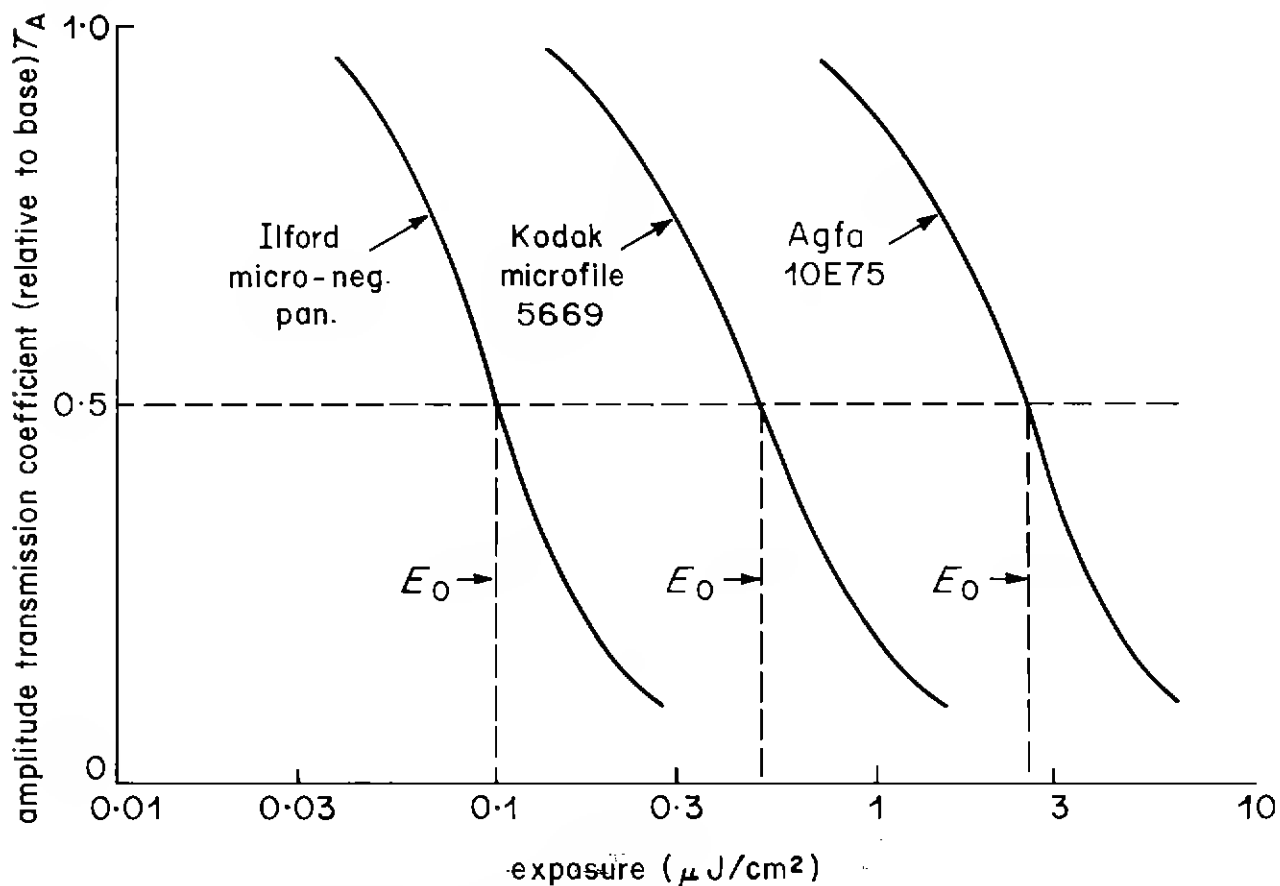
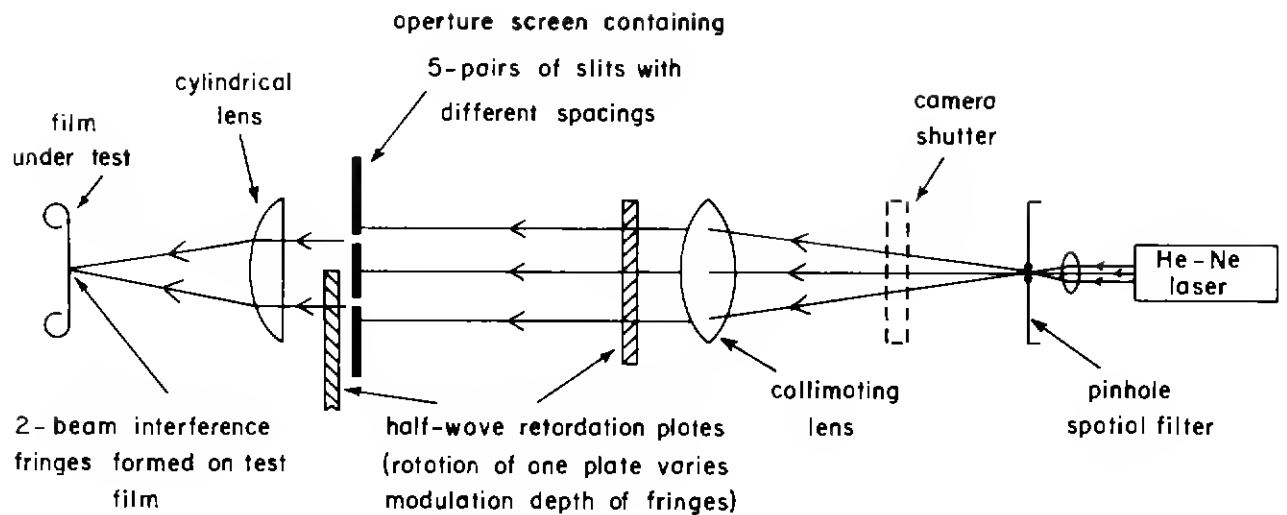
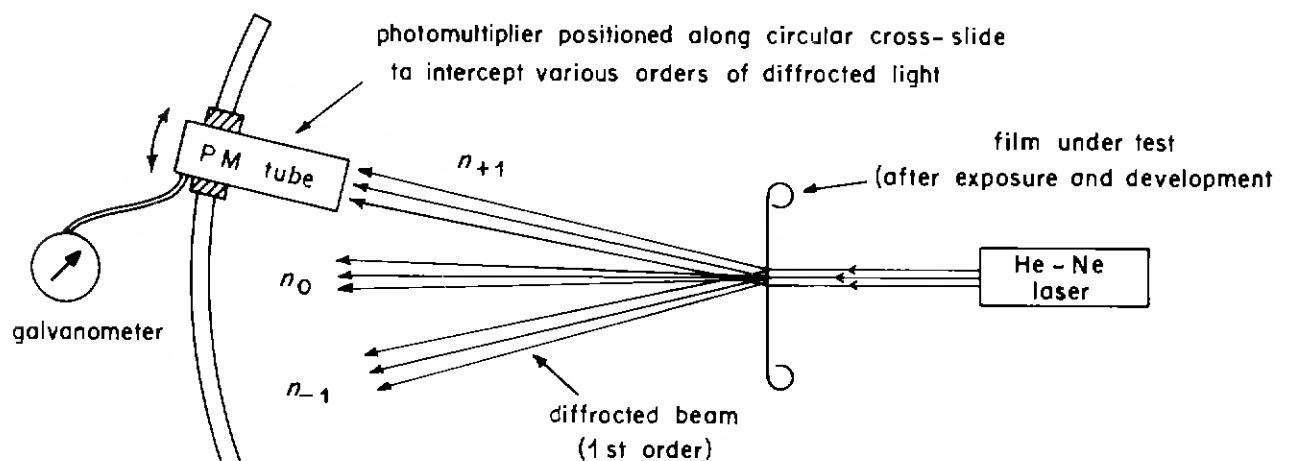


Fig. 11 Transfer characteristics of film emulsions; amplitude transmission T_A versus exposure in $\mu\text{J}/\text{cm}^2$ (log scale)



(a)



(b)

Fig. 12 Arrangement used to measure effective modulation transfer functions of film emulsions up to 200 cycles/mm
a Generation of sinusoidal intensity patterns on film sample
b Measurement of diffraction efficiency

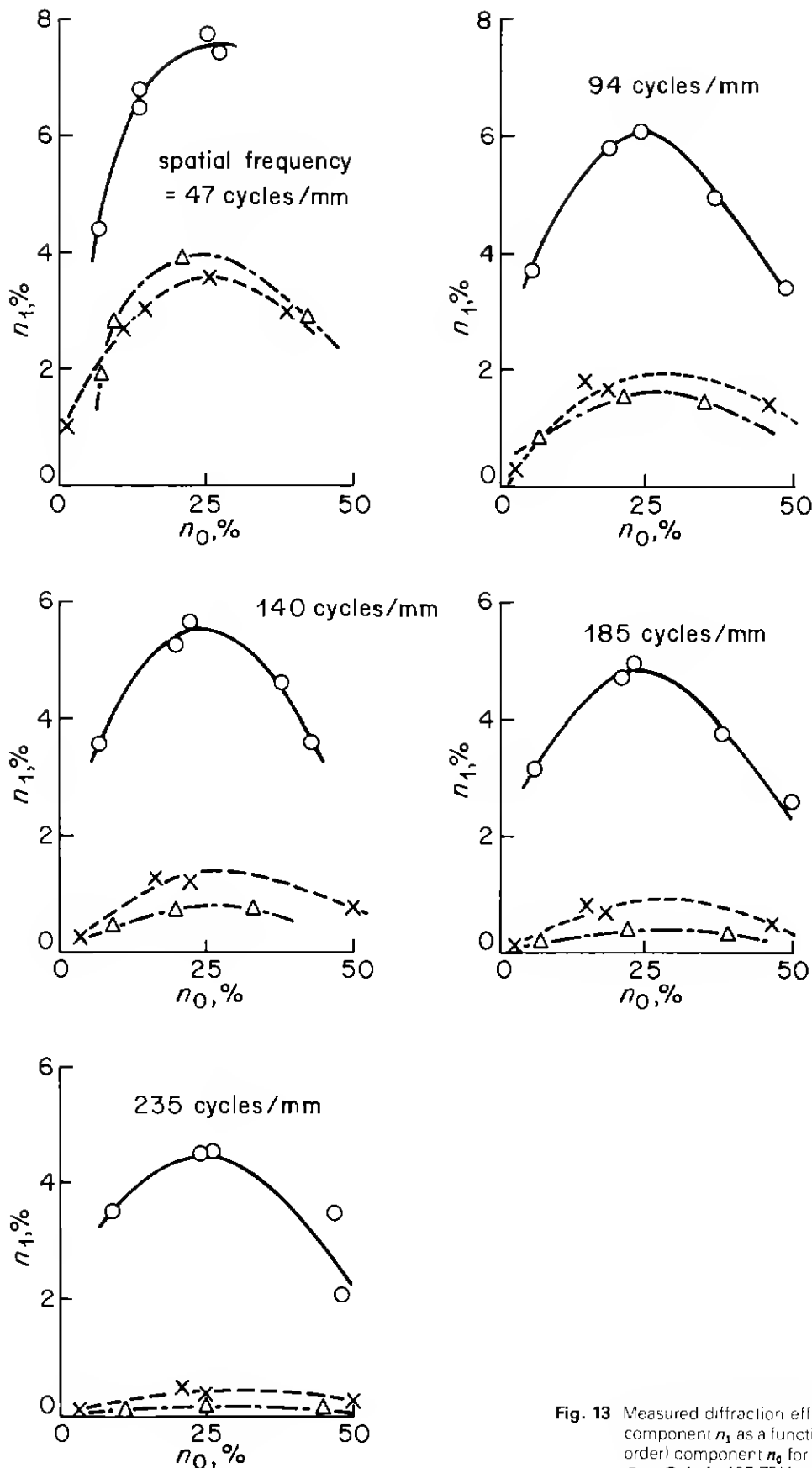


Fig. 13 Measured diffraction efficiency of test samples: first-order component n_1 as a function of the undiffracted (zero order) component n_0 for various spatial frequencies
 ○—○ Agfa 10E 75 ×—× Kodak Microfile 5669
 △—△ Ilford Micro-neg pan

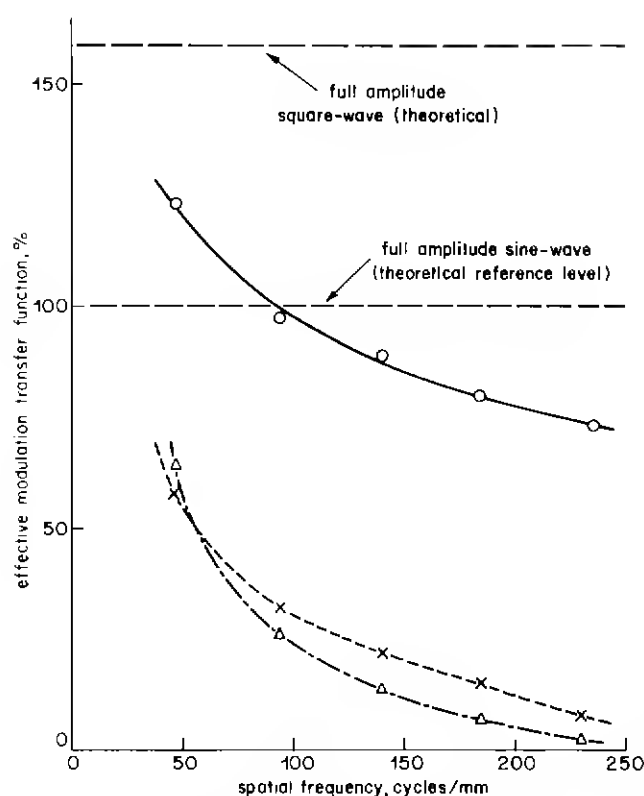


Fig. 14 Effective modulation transfer functions of film samples
 ○—○ Agfa 10E 75 X—X Kodak Microfile 5669
 △—△ Ilford Micro-neg pan

centage of the incident intensity, where the latter is taken to be the transmitted intensity of the incident light in the unexposed region of the record. All results were obtained using He-Ne laser radiation at $\lambda = 632.8\text{nm}$. The zero-order intensity was varied by altering the exposure within the appropriate range of the emulsions' transfer characteristics.

Derived curves, which are equivalent to the MTFs for the films tested, are shown in figure 14, where the maximum first-order diffraction intensity is plotted as a function of spatial frequency for each film type. The measured intensities are shown normalised to the theoretical diffraction intensity for a perfect sine-wave amplitude grating with 100% modulation depth. Also marked on the figure is the theoretical level for the first-order component of a square-wave grating.

It will be noticed that, for one film type (Agfa 10E 75), the diffraction efficiency at low spatial frequencies exceeds the theoretical level for a sine-wave grating. This reflects the fact that the amplitude transmission characteristic of the film has only a limited region of linearity, so that the extreme excursions of exposure occurring with a high modulation depth are 'crushed', resulting in a recorded pattern which tends to have a square-wave form.

There are several other conclusions which can be drawn from these diffraction measurements:

- 1 The optimum diffraction efficiency occurs for a zero-order intensity in the region of 25%. (See figure 13.) This is equivalent to an amplitude transmission coefficient $T_A = 0.5$, which is the region of the transfer

characteristic that is most linear and has the greatest slope (figure 11). This is clearly an optimum bias point for holographic applications where minimum distortion and maximum efficiency are required.

- 2 The range of zero-order intensities for which the diffraction efficiency is greater than 50% of its maximum values, irrespective of spatial frequency, appears to be from 5% to 50% approximately. Taking this range as a practical tolerance for holographic applications, the corresponding latitude for exposure is from $0.3E_0$ to $1.75E_0$, where E_0 is the optimum exposure.
- 3 Although the two most sensitive films have some response at spatial frequencies above 200 cycles/mm, the diffraction efficiency begins to droop severely at much lower spatial frequencies. Their effective MTFs compare most unfavourably with that of the slowest (least sensitive) film.

4 Micro-hologram experiments

An arrangement for exposing film samples to produce low-frequency micro-holograms was constructed as outlined in figure 15a. A fast mechanical shutter was used to obtain exposures in the range 50 to 500 microseconds. The hologram aperture screen and principal lens were mounted on a transverse slide controlled by a screw thread. Hence, by successive lateral displacements and exposures, rows of spot holograms could be produced.

In order to measure the reconstruction properties of the test holograms, an extension to the recording apparatus was constructed as shown in figure 15b. The outer holes in the aperture screen were blanked off and the processed film sample replaced and registered in the film holder. A search probe, consisting of an array of small (1mm dia.) fibre-optic light guides, was placed about 10 cm behind the film holder. The guides were arranged to have the same geometrical disposition as the wanted set of reconstructed images (including the central zero-order components for reference). The relative image intensities were measured by coupling the free end of each guide in turn to a photo-multiplier.

For many of the tests, object-screens were constructed based on the designs described earlier (section 2.4). The collimated light incident on the object-screen was plane-polarised, so that by rotating a piece of half-wave retardation plate (with its axes suitably oriented) over an aperture the polarisation could be varied through 90° . In this condition, no interference can take place in the hologram plane between light from the particular aperture and that from the reference aperture. Hence, the aperture is effectively switched off and it will not appear in the subsequent reconstruction of the hologram.⁸ Moreover, the switching effect is accomplished without altering the average exposure, which would not be the case if the apertures were simply blanked off.

4.1 5-bit holograms (40 μm dia.)

Initially, 5-bit holograms (i.e. 5 outer apertures in the object

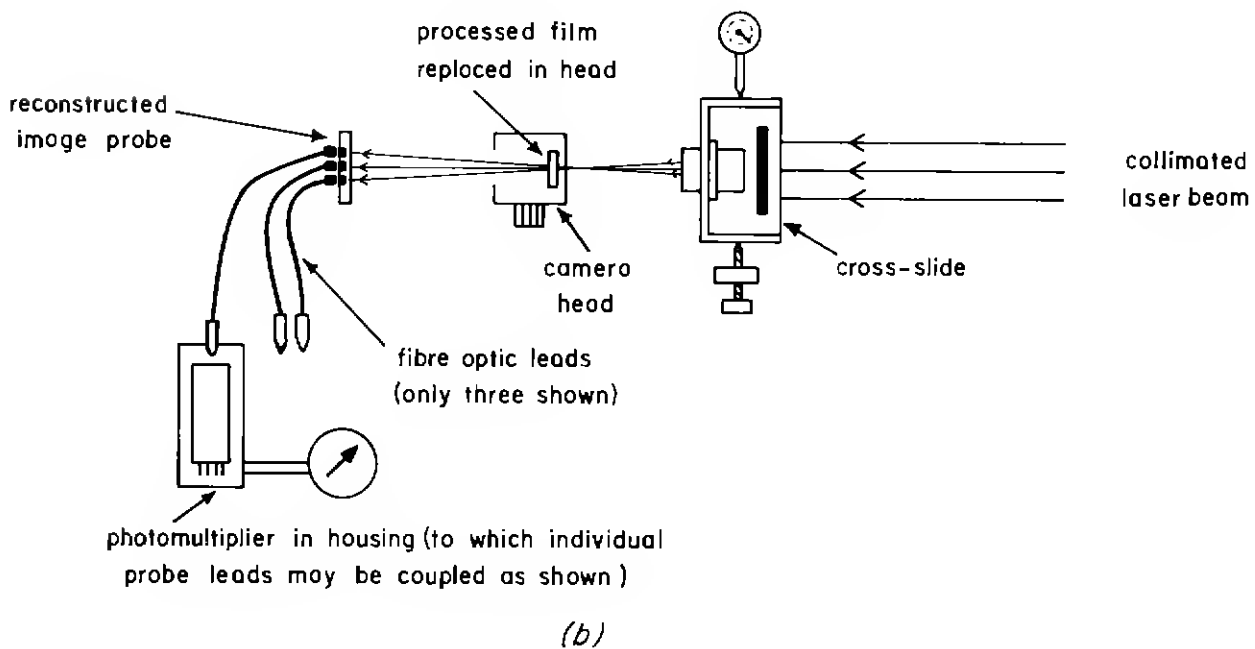
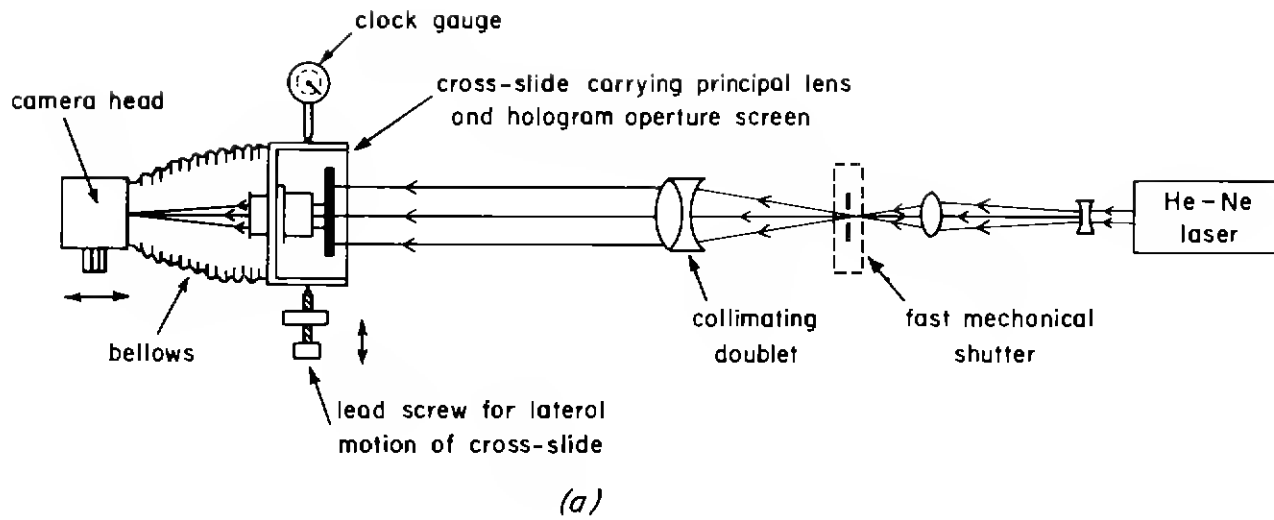
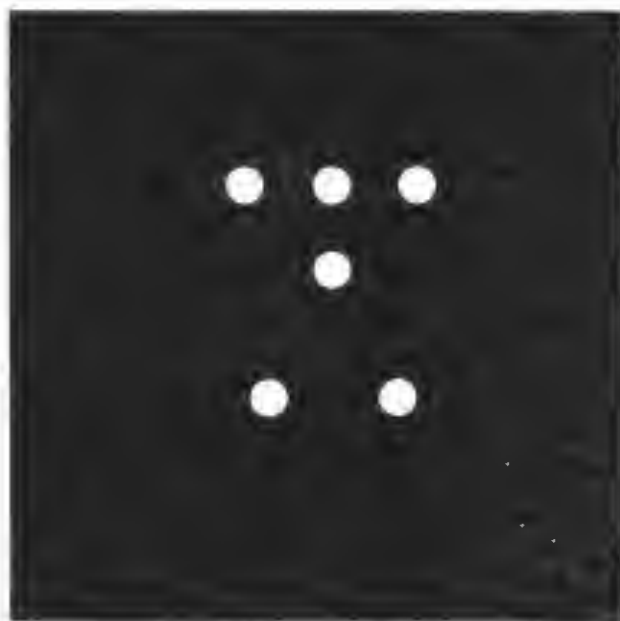
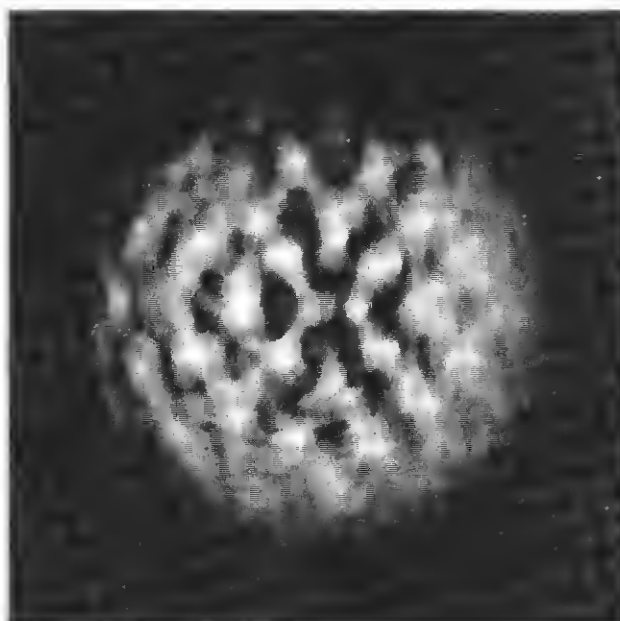


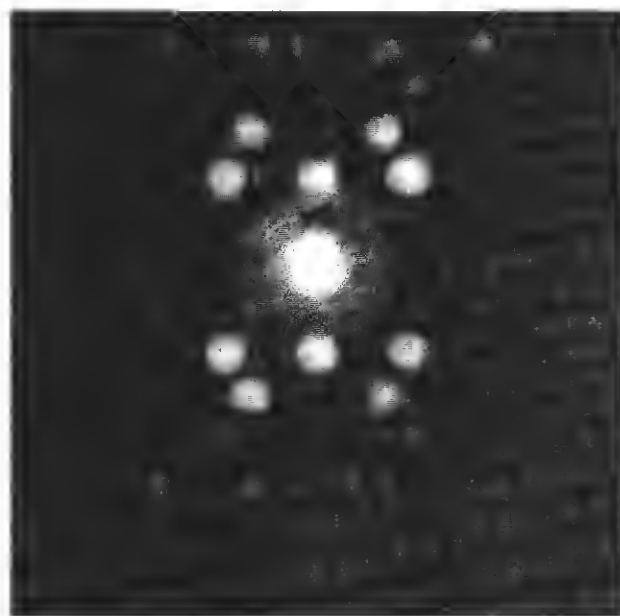
Fig. 15 Arrangements for exposing and measuring micro-holograms
a. Recording the holograms
b. Assessing the reconstructed images



a.



b.



c.

Fig. 16 Formation and reconstruction photographs of a five-bit hologram ($40\mu\text{m}$ dia.)

- a. Transmission pattern of original object screen
- b. The micro-hologram intensity pattern
- c. The reconstructed image pattern (showing image blurring and extraneous images)

screen) were produced having an effective diameter of $40\mu\text{m}$ and a maximum spatial frequency of 120 cycles/mm. Figure 16a is a transmission photograph of the original screen pattern, figure 16b is a photomicrograph of the actual FT hologram while figure 16c is a photograph of the reconstructed image showing the expected image blurring and extraneous images. All but the wanted set of images would normally be masked off.

Similar micro-holograms were formed on each of the film stocks and photo-electric measurements were made of their diffraction efficiency. The measured results are summarised in Table 1 where, for comparison, the last column shows the theoretical maximum expected for each emulsion type.

The next stage was to form a row of 5-bit holograms, spaced $40\mu\text{m}$ apart; starting with no outer apertures 'on' for the first hologram, each subsequent hologram had one more

TABLE 1
5-bit $40\mu\text{m}$ diameter holograms

Emulsion type	MTF at 100 cycles/mm	Hologram Efficiency % (average per bit)	
		Measured	Theoretical (Max.)
Agfa 10E 75	0.98	0.72	1.22
Kodak Microfile 5669	0.3	0.16	0.37
Ilford Micro-Neg Pan	0.24	0.2	0.3

aperture 'open' than its predecessor. During reconstruction, the incident beam was manually scanned over the row of holograms while the corresponding variation of light intensity in each of the (fixed) image locations was determined. These results are shown plotted in figure 17a. With respect to the arbitrary threshold marked on the figure, it will be noted that threshold-to-background intensity ratio is

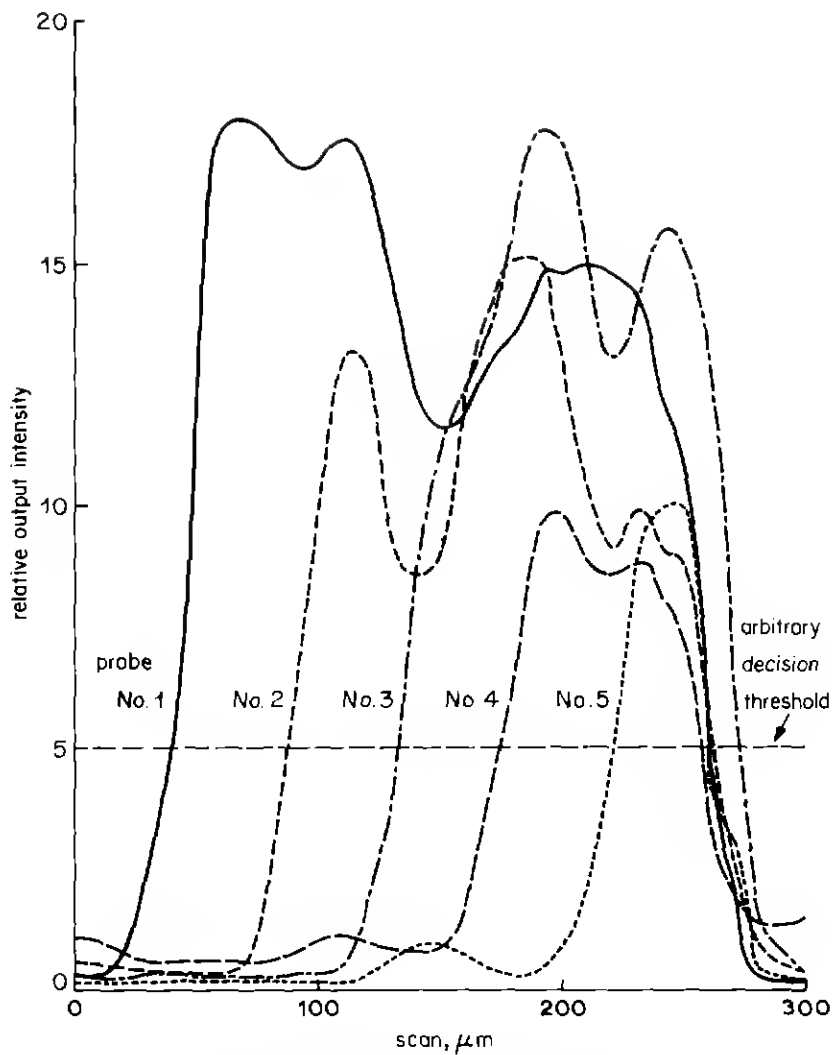
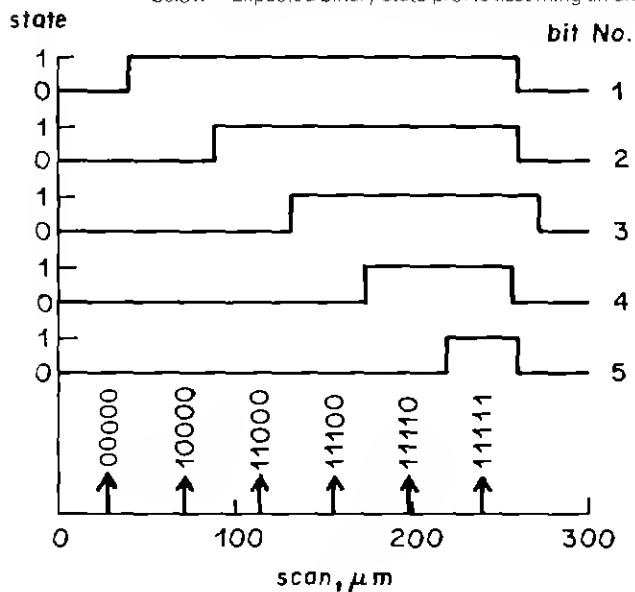


Fig. 17 Reconstruction profiles for a row of five-bit micro-holograms
 above: Measured relative intensities at each of the wanted image positions during a scan of five adjacent holograms
 below: Expected binary state profile assuming an arbitrary decision threshold in Fig. 17a



generally better than 10:1; part of the background light measured was, in fact, due to the veiling glare introduced by the principal lens and is not the result of image crosstalk or noise. If the electrical outputs of five photodetectors optically coupled to each image location respectively, had been connected to simple level comparator circuits, the binary quantised outputs would be approximately those shown in figure 17b.

4.2 9-bit holograms (30 μ m dia.)

The first attempts to form 9-bit holograms with a diameter of 30 μ m, while still restricting the spatial-frequency bandwidth to 120 cycles/mm, were unsuccessful due to inter-bit crosstalk; the object-screen apertures were too close to each other in relation to the diameter of the reference aperture. Better reconstruction imagery was obtained by increasing the separation of the apertures, thus raising the spatial-frequency bandwidth to 160 cycles/mm.

Another difficulty encountered when regularly-spaced apertures occur in lines is that the range of intensities in the exposure of the hologram can greatly exceed the useful range provided by the emulsion transfer characteristics. As the number, N , of screen apertures increases, spiky peaks of intensity proportional to N^2 occur, which are 'crushed' in the recording process and give rise to additional crosstalk in the reconstruction. The intensity range of the exposure can be substantially reduced by deliberately altering the relative phases of the contributions from the different apertures prior to mutual interference in the hologram plane. This can be achieved by defocusing, i.e. by moving the film slightly away from the true Fourier-transform plane, or by altering the path lengths by means of a random-phase plate. The latter appears to be the more satisfactory method for micro-hologram arrays.

Table 2 shows the reconstruction efficiency obtained with

TABLE 2
9-bit, 30 μ m diameter holograms

Emulsion type	MTF at 150 cycles/mm	Hologram Efficiency % (average per bit)	
		Measured	Theoretical (Max.)
Agfa 10E 75	0.85	0.34	0.6

the Agfa 10E 75 emulsion for a 9-bit hologram, 30 μ m in diameter, having a spatial-frequency bandwidth of 160 cycles/mm. As previously, the last column in Table 2 indicates the maximum theoretical efficiency.

4.3 Copying

A potential advantage of data recording using low-spatial-frequency holograms is that copies could be taken using conventional printing techniques. Clearly, emulsions with high resolution characteristics need to be used for the prints. A brief experiment was carried out in which a row of 9-bit holograms, recorded on the Agfa 10E 75 emulsion, was contact printed on the same stock but with modified develop-

ment characteristics; ordinary (tungsten) light was used for the exposure. A similar reconstructed image was obtained with the copied (positive) holograms, but there was a noticeable increase in crosstalk and a decrease in signal-to-background ratio, largely due to the extra blemishes caused by dust and scratches in the process.

More work is required to determine the most suitable copying technique; if an original recording of digital data can be produced with a sufficiently low error rate on read-out, perhaps the best approach would be simply to replay and re-record.

5 Conclusions

The analysis of spot-hologram formation and recording on film, supported by several basic experiments, indicates that storage densities of 10^6 bits/cm² are feasible with average signal-to-interference ratios of 20dB or better in the reconstruction, and with hologram spatial-frequency bandwidths of less than 200 cycles/mm. This storage density would allow a full-bandwidth colour television signal, coded in PCM form (8 bits/sample), to be recorded with the same consumption of medium as that required for a conventional analogue recording of 35mm cine-film. However, only a black-and-white film emulsion is required which, when suitably processed, is attractive for archival storage applications. Greater storage densities are theoretically possible by increasing the spatial-frequency bandwidth, and the holographic advantage of defect resistance is improved by increasing the spot area of each hologram spot in the array. The optimum number of bits per hologram spot, for digital television recording in real time, will depend finally on the development of a suitable scanning method in conjunction with arrays of fast-acting light switches. This latter subject will form the next phase of this feasibility study.

The measured modulation transfer functions of several photographic emulsions used for high-resolution photography and line recording proved to be rather inadequate for micro-holograms with spatial-frequency bandwidths greater than about 100 cycles/mm. Far superior results were obtained using a commercially-available holographic film, but its sensitivity is about one order lower. Even so, for this slower holographic film, it is estimated that real-time recording at a transfer rate of 100 Megabits/sec can be achieved with laser radiation power less than 10mW. This figure assumes a packing density of 10^6 bits/cm² or greater and an optical efficiency of the recording apparatus better than 10%. A similar order of laser power is required for the reconstruction process. Hence low-power helium-neon gas lasers would be suitable light sources, and these are now available at relatively low cost and with a tube life in excess of 5000 hours.

The experimental work shows that the exposure latitude for recording micro-holograms is rather small ($\pm\frac{1}{2}$ stop) and this may present some operational difficulties bearing in mind possible batch-to-batch variations in film sensitivity. It was envisaged that, by recording micro-hologram arrays

with a conventional aspect ratio, a faint analogue version of the picture could be simultaneously overlaid on the digital recording; this additional intensity variation could occupy most, if not all, of the exposure latitude and may not be practicable.

6 References

1. Rajchman, J.A. Holographic optical memory: an optical read-write mass memory. *Appl. Optics*, 1970, Vol. 9, 10, pp. 2269-2271.
2. Mezrich, R.S. Magnetic holography. *Appl. Optics*, 1970, Vol. 9, 10, pp. 2275-2279.
3. Lin, L.H. and Beauchamp, H.L. Write-read-erase in situ optical memory using thermoplastic holograms. *Appl. Optics*, 1970, Vol. 9, 9, pp. 2088-2092.
4. Lunn, J.D., Moffat, M.E.B. Possible techniques for the recording of digital television signals. BBC Research Department Report No. 1969/42.
5. Briggs, D.A., Hachfeld, K.D., Reid, D.C.J. and Waterworth, P. A holographic data storage system. *Systems Technology*, 1970, 9, pp. 29-32.
6. Leith, E. and Upatnieks, J. Wavefront reconstruction with diffused illumination and three-dimensional objects. *J. Opt. Soc. Am.*, 1964, Vol. 54, 11, p. 1295.
7. Taylor, E.W. Holography: General Principles and Applications to Sound and Television Recording. Article in *this issue*, p. 4.
8. British Patent Application No. 7954/71.
9. Blackmer, L.L., Vankerkhove, A.P., Baldwin, R.E. and Stultz, K.F. Some aspects of multiple-beam interference techniques in digital data recording. *Appl. Optics*, 1970, Vol. 9, 12, pp. 2753-2761.
10. Born, M. and Wolf, E. *Principles of optics*. London, Pergamon, 1959.
11. Bestenriener, R., Greis, U. and Weiershausen, W. Alpha-numeric storage capacity of a defocused Fourier hologram matrix. *Phot. Science and Engineering*, 1972, Vol. 16, 1, pp. 4-15.

Contributors to this issue



E. W. Taylor is a graduate of Emmanuel College, Cambridge and joined the BBC as a Graduate Apprentice in 1954 after spending two years with the Radio Group of the Cavendish Laboratory. He joined the Research Department in 1956 and was involved until 1964 with field trials of UHF television transmissions and the determination of co-channel interference limits. Since 1964 he has been concerned with television signal origination and display, particularly in colorimetric and other optical fields. Recently he has been working on problems arising from the use of metal halide discharge lamps for film and television scene lighting.



Kenneth Hacking joined the BBC Research Department in 1955, after working in an industrial research laboratory mainly concerned with thin films and vacuum deposition techniques. The early part of his career in Research Department was with Optics Section dealing with the performance of lenses for television, assessment of image quality, and colorimetric analysis in colour television systems.

He was transferred in 1965 to Transmission Section, Radio Group, and worked on prediction problems associated with the propagation of UHF/VHF signals over hills. Later he moved to Radio Frequency Systems Section for a short period to study the possible application of less conventional modulation systems for m.f. sound broadcasting.

At present he is with the Physics Section of Studio Group.